

Année 2024

2024 TOU3 1130
2024 TOU3 1131

THESE
POUR LE DIPLÔME D'ÉTAT DE DOCTEUR EN MÉDECINE
SPECIALITÉ MÉDECINE GÉNÉRALE

Présentée et soutenue publiquement
par

Élise AGUERA
Et
Cécile FERRÉ

Le 17 Octobre 2024

**REVUE SYSTÉMATIQUE DE LA LITTÉRATURE SUR L'UTILISATION DE L'INTELLIGENCE
ARTIFICIELLE DANS LE DIAGNOSTIC DES MÉLANOMES ET DES CARCINOMES NON
MÉLANOCYTAIRES EN SOINS PRIMAIRES**

Directeur de thèse : Pr Jordan BIREBENT

JURY :

Madame le Professeur Julie DUPOUY

Présidente

Monsieur le Professeur Jean-Christophe POUTRAIN

Assesseur

Monsieur le Professeur Jordan BIREBENT

Assesseur



**UNIVERSITÉ
TOULOUSE III
PAUL SABATIER**

FACULTÉ DE SANTÉ

Département de Médecine, Maïeutique et Paramédical

Doyen – Directeur : Pr Thomas GEERAERTS

Tableau du personnel Hospitalo-Universitaire de médecine

2023-2024

Professeurs Honoraires

Doyen Honoraire	M. CHAP Hugues	Professeur Honoraire	M. GERAUD Gilles
Doyen Honoraire	M. GUIRAUD-CHAUMEIL Bernard	Professeur Honoraire	M. GHISOLFI Jacques
Doyen Honoraire	M. PUEL Pierre	Professeur Honoraire	M. GLOCK Yves
Doyen Honoraire	M. ROUGE Daniel	Professeur Honoraire	M. GOUZI Jean-Louis
Doyen Honoraire	M. SERRANO Ele	Professeur Honoraire	M. GRAND Alain
Doyen Honoraire	M. VINEL Jean-Pierre	Professeur Honoraire	M. HOFF Jean
Professeur Honoraire	M. ABBAL Michel	Professeur Honoraire	M. JOFFRE Francis
Professeur Honoraire	M. ADER Jean-Louis	Professeur Honoraire	M. LAGARRIGUE Jacques
Professeur Honoraire	M. ADDUE Daniel	Professeur Honoraire	M. LANG Thierry
Professeur Honoraire	M. ARBUS Louis	Professeur Honoraire	Mme LARENG Marie-Blanche
Professeur Honoraire	M. ARLET Philippe	Professeur Honoraire	M. LAROCHE Michel
Professeur Honoraire	M. ARLET-SUAU Elisabeth	Professeur Honoraire	M. LAUQUE Dominique
Professeur Honoraire	M. ARNE Jean-Louis	Professeur Honoraire	M. LAURENT Guy
Professeur Honoraire	M. ATTAL Michel	Professeur Honoraire	M. LAZORTHES Franck
Professeur Honoraire	M. BARRET André	Professeur Honoraire	M. LEOPHONTE Paul
Professeur Honoraire	M. BARTHE Philippe	Professeur Honoraire	M. MAGNAVAL Jean-François
Professeur Honoraire	M. BAYARD Francis	Professeur Honoraire	M. MALECAZE François
Professeur Honoraire	M. BLANCHER Antoine	Professeur Honoraire	M. MANELFE Claude
Professeur Honoraire	M. BOCCALON Henri	Professeur Honoraire	M. MANSAT Michel
Professeur Honoraire	M. BONAFÉ Jean-Louis	Professeur Honoraire	M. MARCHOU Bruno
Professeur Honoraire	M. BONEU Bernard	Professeur Honoraire	M. MASSIP Patrice
Professeur Honoraire	M. BONNEVILLE Paul	Professeur Honoraire	Mme MARTY Nicole
Professeur Honoraire	M. BOSSAVY Jean-Pierre	Professeur Honoraire	M. MAZIERES Bernard
Professeur Honoraire	M. BOUNHCURE Jean-Paul	Professeur Honoraire	M. MONROZIES Xavier
Professeur Honoraire	M. BOUTAULT Franck	Professeur Honoraire	M. MONTASTRUC Jean-Louis
Professeur Honoraire Associé	M. BROS Bernard	Professeur Honoraire	M. MOSCOVICI Jacques
Professeur Honoraire	M. BUGAT Roland	Professeur Honoraire	M. MURAT
Professeur Honoraire	M. BUJAN Louis	Professeur Honoraire associé	M. NICODEME Robert
Professeur Honoraire	M. CAHUZAC Jean-Philippe	Professeur Honoraire	M. OLIVES Jean-Pierre
Professeur Honoraire	M. CALVAS Patrick	Professeur Honoraire	M. PARINAUD Jean
Professeur Honoraire	M. CARATERO Claude	Professeur Honoraire	M. PASCAL Jean-Pierre
Professeur Honoraire	M. CARLES Pierre	Professeur Honoraire	M. PERRET Bertrand
Professeur Honoraire	M. CARON Philippe	Professeur Honoraire	M. PESSEY Jean-Jacques
Professeur Honoraire	M. CARRIERE Jean-Paul	Professeur Honoraire	M. PLANTE Pierre
Professeur Honoraire	M. CARTON Michel	Professeur Honoraire	M. PONTONNIER Georges
Professeur Honoraire	M. CATHALA Bernard	Professeur Honoraire	M. POURRAT Jacques
Professeur Honoraire	M. CHABANON Gérard	Professeur Honoraire	M. PRADERE Bernard
Professeur Honoraire	M. CHAMONTIN Bernard	Professeur Honoraire	M. PRIS Jacques
Professeur Honoraire	M. CHAVOIN Jean-Pierre	Professeur Honoraire	Mme PUEL Jacqueline
Professeur Honoraire	M. CHIRON Philippe	Professeur Honoraire	M. PUJOL Michel
Professeur Honoraire	M. CLANET Michel	Professeur Honoraire	M. QUERLEU Denis
Professeur Honoraire	M. CONTE Jean	Professeur Honoraire	M. RAILHAC Jean-Jacques
Professeur Honoraire	M. COSTAGLICLA Michel	Professeur Honoraire	M. REGNIER Claude
Professeur Honoraire	M. COTONAT Jean	Professeur Honoraire	M. REME Jean-Michel
Professeur Honoraire	M. DABERNAT Henri	Professeur Honoraire	M. RISCHMANN Pascal
Professeur Honoraire	M. DAHAN Marcel	Professeur Honoraire	M. RIVIERE Daniel
Professeur Honoraire	M. DALOUS Antoine	Professeur Honoraire	M. ROCHE Henri
Professeur Honoraire	M. DALY-SCHWEITZER Nicolas	Professeur Honoraire	M. ROCHICCIOLI Pierre
Professeur Honoraire	M. DAVID Jean-Frédéric	Professeur Honoraire	M. ROLLAND Michel
Professeur Honoraire	Mme DELISLE Marie-Bernadette	Professeur Honoraire	M. ROQUES-LATRILLE Christian
Professeur Honoraire	M. DELSOL Georges	Professeur Honoraire	M. ROUGE Daniel
Professeur Honoraire	Mme DIDIER Jacqueline	Professeur Honoraire	M. RUMEAU Jean-Louis
Professeur Honoraire	M. DUCOS Jean	Professeur Honoraire	M. SALVADOR Michel
Professeur Honoraire	M. DUFFAUT Michel	Professeur Honoraire	M. SALVAYRE Robert
Professeur Honoraire	M. DUPRE M.	Professeur Honoraire	M. SARRAMON Jean-Pierre
Professeur Honoraire	M. DURAND Dominique	Professeur Honoraire	M. SCHMITT Laurent
Professeur Honoraire associé	M. DUTAU Guy	Professeur Honoraire	M. SERRE Guy
Professeur Honoraire	M. ESCOURROU Jean	Professeur Honoraire	M. SIZUN Jacques
Professeur Honoraire	M. ESQUERRE Jean-Paul	Professeur Honoraire	M. SIMON Jacques
Professeur Honoraire	M. FABIE Michel	Professeur Honoraire	M. SUC Jean-Michel
Professeur Honoraire	M. FABRE Jean	Professeur Honoraire	M. THOUVENOT Jean-Paul
Professeur Honoraire	M. FOURNIAL Gérard	Professeur Honoraire	M. TREMOULET Michel
Professeur Honoraire	M. FOURNIE Bernard	Professeur Honoraire	M. VALDIGUIE Pierre
Professeur Honoraire	M. FORTANIER Gilles	Professeur Honoraire	M. VAYSSE Philippe
Professeur Honoraire	M. FRAYSSE Bernard	Professeur Honoraire	M. VIRENQUE Christian
Professeur Honoraire	M. FREXINOS Jacques	Professeur Honoraire	M. VOIGT Jean-Jacques
Professeur Honoraire	Mme GENESTAL Michèle		

Professeurs Émérites

Professeur BUJAN Louis
 Professeur CARON Philippe
 Professeur CHAP Hugues
 Professeur FRAYSSE Bernard
 Professeur LANG Thierry

Professeur LAROCHE Michel
 Professeur LAUQUE Dominique
 Professeur MAGNAVAL Jean-François
 Professeur MARCHOU Bruno
 Professeur MESTHE Pierre

Professeur MONTASTRUC Jean-Louis
 Professeur PARINI Angelo
 Professeur PERRET Bertrand
 Professeur ROQUES LATRILLE Christian
 Professeur SERRE Guy

Professeur SIZUN Jacques
 Professeur VIRENQUE Christian
 Professeur VINEL Jean-Pierre

FACULTÉ DE SANTÉ
Département de Médecine, Maïeutique et Paramédical

P.U. - P.H.
Classe Exceptionnelle et 1ère classe

M. ACAR Philippe	Pédiatrie	Mme LAFRIE Anne	Radiothérapie
M. ACCADBLED Franck (C.E)	Chirurgie infantile	M. LARRUE Vincent	Neurologie
M. ALRIC Laurent (C.E)	Médecine Interne	M. LAUQUE Dominique (C.E)	Médecine d'Urgence
M. AMAR Jacques (C.E)	Thérapeutique	Mme LAURENT Camille	Anatomie Pathologique
Mme ANDRIEU Sandrine	Epidémiologie, Santé publique	M. LAUWERS Frédéric	Chirurgie maxillo-faciale
M. ARBUS Christophe	Psychiatrie	M. LE CAIGNEC Cédric	Génétique
M. ARNAL Jean-François (C.E)	Physiologie	M. LEVADE Thierry (C.E)	Biochimie
M. AUSSEIL Jérôme	Biochimie et biologie moléculaire	M. LIBLAU Roland (C.E)	Immunologie
M. AVET-LOISEAU Hervé (C.E)	Hématologie, transfusion	M. MALAVALD Bernard (C.E)	Urologie
M. BERRY Antoine (C.E.)	Parasitologie	M. MANSAT Pierre (C.E)	Chirurgie Orthopédique
Mme BERRY Isabelle (C.E)	Biophysique	M. MARCHEIX Bertrand	Chirurgie thoracique cardiovasculaire
M. BIRMES Philippe	Psychiatrie	M. MARQUE Philippe (C.E)	Médecine Physique et Réadaptation
M. BONNEVILLE Nicolas	Chirurgie orthopédique et traumatologique	M. MARTIN-BONDEL Guillaume	Maladies infectieuses, maladies tropicales
M. BONNEVILLE Fabrice	Radiologie	M. MAS Emmanuel	Pédiatrie
M. BROUCHET Laurent	Chirurgie thoracique et cardio-vasculaire	M. MAURY Jean-Philippe (C.E)	Cardiologie
M. BROUSSET Pierre (C.E)	Anatomie pathologique	Mme MAZEREEUW Juliette	Dermatologie
Mme BURA-RIVIERE Alessandra (C.E)	Médecine Vasculaire	M. MAZIERES Julien (C.E)	Pneumologie
M. BUREAU Christophe (C.E.)	Hépatogastro-entérologie	M. MINVILLE Vincent (C.E.)	Anesthésiologie Réanimation
M. BUSCAIL Louis (C.E.)	Hépatogastro-entérologie	M. MOLINIER Laurent (C.E)	Epidémiologie, Santé Publique
M. CANTAGREL Alain (C.E)	Rhumatologie	Mme MOYAL Elisabeth (C.E)	Cancérologie
M. CARRERE Nicolas	Chirurgie Générale	M. MUSCARI Fabrice	Chirurgie Digestive
M. CARRIE Didier (C.E)	Cardiologie	Mme NOURHASHEMI Fatemeh (C.E)	Gériatrie
M. CHAIX Yves	Pédiatrie	M. OLVOT Jean-Marc	Neurologie
Mme CHANTALAT Eklodie	Anatomie	M. OSWALD Eric (C.E)	Bactériologie-Virologie
M. CHAPUT Benoit	Chirurgie plastique	M. PAGES Jean-Christophe	Biologie cellulaire
Mme CHARPENTIER Sandrine (C.E)	Médecine d'urgence	M. PARIENTE Jérémie	Neurologie
M. CHAUFOUR Xavier (C.E.)	Chirurgie Vasculaire	M. PAUL Carle (C.E)	Dermatologie
M. CHAUVÉAU Dominique	Néphrologie	M. PAYOUX Pierre (C.E)	Biophysique
M. CHAYNES Patrick	Anatomie	M. PAYRASTRE Bernard (C.E)	Hématologie
M. CHOLLET François (C.E)	Neurologie	M. PERON Jean-Marie (C.E)	Hépatogastro-entérologie
M. CONSTANTIN Amaud	Rhumatologie	Mme PERROT Aurone	Physiologie
M. COURBON Frédéric (C.E)	Biophysique	M. RASCOL Olivier (C.E)	Pharmacologie
Mme COURTADE SAÏDI Monique (C.E)	Histologie Embryologie	Mme RAUZY Odile (C.E.)	Médecine Interne
M. DAMBRIN Camille	Chir. Thoracique et Cardiovasculaire	M. RAYNAUD Jean-Philippe (C.E)	Psychiatrie infantile
M. DE BOISSEZON Xavier	Médecine Physique et Réadapt Fonct.	M. RECHER Christian(C.E)	Hématologie
M. DEGUINE Olivier (C.E)	Oto-rhino-laryngologie	M. RITZ Patrick (C.E)	Nutrition
M. DELABESSE Eric	Hématologie	M. ROLLAND Yves (C.E)	Gériatrie
M. DELOBEL Pierre	Maladies Infectieuses	M. RONCALLI Jérôme	Cardiologie
M. DELORD Jean-Pierre (C.E)	Cancérologie	M. ROUSSEAU Hervé (C.E)	Radiologie
M. DIDIER Alain (C.E)	Pneumologie	M. ROUX Franck-Emmanuel (C.E.)	Neurochirurgie
M. DUCCOMMUN Bernard	Cancérologie	M. SAILLER Laurent (C.E)	Médecine Interne
Mme DULY-BOUHANICK Béatrice (C.E)	Thérapeutique	M. SALES DE GAUZY Jérôme (C.E)	Chirurgie Infantile
M. ELBAZ Meyer	Cardiologie	M. SALLES Jean-Pierre (C.E)	Pédiatrie
Mme EVRARD Solène	Histologie, embryologie et cytologie	M. SANS Nicolas	Radiologie
M. FERRIERES Jean (C.E)	Epidémiologie, Santé Publique	Mme SAVAGNER Frédérique	Biochimie et biologie moléculaire
M. FOURCADE Olivier (C.E)	Anesthésiologie	Mme SELVES Janick (C.E)	Anatomie et cytologie pathologiques
M. FOURNIÉ Pierre	Ophthalmologie	M. SENARD Jean-Michel (C.E)	Pharmacologie
M. GALINIER Michel (C.E)	Cardiologie	M. SERRANO Elie (C.E)	Oto-rhino-laryngologie
M. GAME Xavier (C.E)	Urologie	M. SOL Jean-Christophe	Neurochirurgie
Mme GARDETTE Virginie	Epidémiologie, Santé publique	M. SOLER Vincent	Ophthalmologie
Mme GASCOIN Géraldine	Pédiatrie	Mme SOMMET Agnès	Pharmacologie
M. GEERAERTS Thomas	Anesthésiologie et réanimation	Mme SOTO-MARTIN Maria-Espina	Gériatrie et biologie du vieillissement
Mme GOMEZ-BROUCHET Anne-Mariel (C.E)	Anatomie Pathologique	M. SOULAT Jean-Marc (C.E)	Médecine du Travail
M. GOURDY Pierre (C.E)	Endocrinologie	M. SOULIE Michel (C.E)	Urologie
M. GRILLEAU RAOUX Jean-Louis (C.E)	Chirurgie plastique	M. SUC Bertrand	Chirurgie Digestive
Mme GUIMBAUD Rosine	Cancérologie	Mme TAUBER Marie-Thérèse (C.E)	Pédiatrie
Mme HANAIRE Hélène (C.E)	Endocrinologie	M. TELMON Norbert (C.E)	Médecine Légale
M. HUYGHE Eric	Urologie	Mme TREMOLLIERES Florence (C.E.)	Biologie du développement
M. IZOPET Jacques (C.E)	Bactériologie-Virologie	Mme URO-COSTE Emmanuelle (C.E)	Anatomie Pathologique
M. KAMAR Nassim (C.E)	Néphrologie	M. VAYSSIERE Christophe (C.E)	Gynécologie Obstétrique
Mme LAMANT Laurence (C.E)	Anatomie Pathologique	M. VELLAS Bruno (C.E)	Gériatrie
M. LANGIN Dominique (C.E)	Nutrition	M. VERGEZ Sébastien	Oto-rhino-laryngologie

P.U. Médecine générale

Mme DUPOUY Julie
M. OUSTRIC Stéphane (C.E)
Mme ROUGE-BUGAT Marie-Eve

FACULTÉ DE SANTÉ
Département de Médecine, Maïeutique et Paramédical

P.U. - P.H.
2ème classe

M. ABBO Olivier	Chirurgie infantile
Mme BONGARD Vanina	Epidémiologie, Santé publique
M. BOUNES Vincent	Médecine d'urgence
Mme BOURNET Barbara	Gastro-entérologie
Mme CASPER Charlotte	Pédiatrie
M. CAVIGNAC Etienne	Chirurgie orthopédique et traumatologie
M. COGNARD Christophe	Radiologie
Mme CORRE Jil	Hématologie
Mme DALENG Florence	Cancérologie
M. DE BONNECAZE Guillaume	Anatomie
M. DECRAMER Stéphane	Pédiatrie
Mme DUPRET-BORIES Agnès	Oto-rhino-laryngologie
M. EDOUARD Thomas	Pédiatrie
M. FAGUER Stanislas	Néphrologie
Mme FARUCH BILFELD Marie	Radiologie et imagerie médicale
M. FRANCHITTO Nicolas	Addictologie
M. GARRIDO-STÖWHAS Ignacio	Chirurgie Plastique
M. GUERBY Paul	Gynécologie-Obstétrique
M. GUIBERT Nicolas	Pneumologie
M. GUILLEMINAULT Laurent	Pneumologie
M. HOUZE-CERFON	Médecine d'urgence
M. HERIN Fabrice	Médecine et santé au travail
M. LAIREZ Olivier	Biophysique et médecine nucléaire
M. LEANDRI Roger	Biologie du dével. et de la reproduction
M. LHERMUSIER Thibault	Cardiologie
M. LOPEZ Raphaël	Anatomie
Mme MARTINEZ Alejandra	Gynécologie
M. MARX Mathieu	Oto-rhino-laryngologie
M. MEYER Nicolas	Dermatologie
Mme MOKRANE Fatma	Radiologie et imagerie médicale
Mme MONTASTIER Emilie	Nutrition
Mme PASQUET Mariène	Pédiatrie
M. PIAU Antoine	Médecine interne
M. PORTIER Guillaume	Chirurgie Digestive
M. PUGNET Grégory	Médecine interne
M. REINA Nicolas	Chirurgie orthopédique et traumatologique
M. RENAUDINEAU Yves	Immunologie
M. REVET Alexis	Pédo-psychiatrie
M. ROUMIGUIE Mathieu	Urologie
Mme RUYSSSEN-WITRAND Adeline	Rhumatologie
M. SAVALL Frédéric	Médecine légale
M. SILVA SIFONTES Stein	Réanimation
M. TACK Ivan	Physiologie
Mme VAYSSE Charlotte	Cancérologie
Mme VEZZOSI Delphine	Endocrinologie
M. YRONDI Antoine	Psychiatrie
M. YSEBAERT Luc	Hématologie

Professeurs Associés

Professeurs Associés de Médecine Générale

M. ABITTEBOUL Yves
M. BIREBENT Jordan
M. BOYER Pierre
Mme FREYENS Anne
Mme IRI-DELAHAYE Motoko
Mme LATROUS Lelia
M. POUTRAIN Jean-Christophe
M. STILLMUNKES André

Professeurs Associés Honoraires

Mme MALAUAUD Sandra
Mme PAVY LE TRACON Anne
M. SIBAUD Vincent
Mme WOISARD Virginie

FACULTÉ DE SANTÉ
Département de Médecine, Maïeutique et Paramédical

MCU - PH

Mme ABRAVANEL Florence	Bactériologie Virologie Hygiène	M. GASQ David	Physiologie
M. APOIL Pol Andre	Immunologie	M. GATIMEL Nicolas	Médecine de la reproduction
Mme ARNAUD Catherine	Epidémiologie	Mme GENNERO Isabelle	Biochimie
Mme AUSSEIL-TRUDEL Stéphanie	Biochimie	Mme GENOUX Annelise	Biochimie et biologie moléculaire
Mme BASSET Céline	Cytologie et histologie	Mme GRARE Marion	Bactériologie Virologie Hygiène
Mme BELLIERES-FABRE Julie	Néphrologie	Mme GUILBEAU-FRUGIER Céline	Médecine légale et droit de la santé
Mme BENEVENT Justine	Pharmacologie fondamentale	Mme GUYONNET Sophie	Nutrition
Mme BERTOLI Sarah	Hématologie, transfusion	M. HAMDJ Safouane	Biochimie
M. BIETH Eric	Génétique	Mme HITZEL Anne	Biophysique
Mme BOST Chloé	Immunologie	M. HOSTALRICH Aurélien	Chirurgie vasculaire
Mme BOUNES Fanny	Anesthésie-Réanimation	M. IRIART Xavier	Parasitologie et mycologie
Mme BREHIN Camille	Pneumologie	Mme JONCA Nathalie	Biologie cellulaire
M. BUSCAIL Etienne	Chirurgie viscérale et digestive	M. KARSENTY Clément	Cardiologie
Mme CAMARE Caroline	Biochimie et biologie moléculaire	M. LAPEBIE François-Xavier	Médecine vasculaire
Mme CANTERO Anne-Valérie	Biochimie	Mme LAPEYRE-MESTRE Maryse	Pharmacologie
Mme CARFAGNA Luana	Pédiatrie	M. LEPAGE Benoit	Biostatistiques et Informatique médicale
Mme CASPAR BAUGUIL Sylvie	Nutrition	M. LHOMME Sébastien	Bactériologie-virologie
Mme CASSAGNE Myriam	Ophthalmologie	Mme MASSIP Clémence	Bactériologie-virologie
Mme CASSAING Sophie	Parasitologie	Mme MAULAT Charlotte	Chirurgie digestive
Mme CASSOL Emmanuelle	Biophysique	Mme MAUPAS SCHWALM Françoise	Biochimie
M. CHASSAING Nicolas	Génétique	M. MONTASTRUC François	Pharmacologie
M. CLAVEL Cyril	Biologie Cellulaire	Mme MOREAU Jessika	Biologie du dév. Et de la reproduction
Mme COLOMBAT Magali	Anatomie et cytologie pathologiques	Mme MOREAU Marion	Physiologie
M. COMONT Thibault	Médecine Interne	M. MOULIS Guillaume	Médecine Interne
M. CONGY Nicolas	Immunologie	Mme NOGUEIRA Maria Léonor	Biologie Cellulaire
Mme COURBON Christine	Pharmacologie	Mme PERICART Sarah	Anatomie et cytologie pathologiques
M. CUROT Jonathan	Neurologie	M. PILLARD Fabien	Physiologie
Mme DAMASE Christine	Pharmacologie	Mme PLAISANCIE Julie	Génétique
Mme DE GLISEZINSKY Isabelle	Physiologie	Mme PUISSANT Bénédicte	Immunologie
M. DEDOUIT Fabrice	Médecine Légale	Mme QUELVEN Isabelle	Biophysique et médecine nucléaire
M. DEGBOE Yannick	Rhumatologie	Mme RAYMOND Stéphanie	Bactériologie Virologie Hygiène
M. DELMAS Clément	Cardiologie	Mme RIBES-MAUREL Agnès	Hématologie
M. DELPLA Pierre-André	Médecine Légale	Mme SABOURDY Frédérique	Biochimie
M. DESPAS Fabien	Pharmacologie	Mme SALLES Juliette	Psychiatrie adultes/Addictologie
M. DUBOIS Damien	Bactériologie Virologie Hygiène	Mme SAUNE Karine	Bactériologie Virologie
Mme ESQUIROL Yolande	Médecine du travail	Mme SIEGFRIED Aurore	Anatomie et cytologie pathologiques
Mme FABBRI Margherita	Neurologie	Mme TRAMUNT Blandine	Endocrinologie, diabète
Mme FILLAUX Judith	Parasitologie	M. TREINER Emmanuel	Immunologie
Mme FLOCH Pauline	Bactériologie-Virologie	Mme VALLET Marion	Physiologie
Mme GALINIER Anne	Nutrition	M. VERGEZ François	Hématologie
M. GANTET Pierre	Biophysique	Mme VIJA Lavinia	Biophysique et médecine nucléaire

M.C.U. Médecine générale

M. BRILLAC Thierry
M. CHICOLLAA Bruno
M. ESCOURROU Emile
Mme GIMENEZ Laetitia

Maîtres de Conférence Associés

M.C.A. Médecine Générale

Mme BOURGEOIS Odile
Mme BOUSSIER Nathalie
Mme DURRIEU Florence
Mme FRANZIN Emile
M. GACHIES Hervé
M. PEREZ Denis
M. PIPONNIER David
Mme PUECH Marielle
M. SAVIGNAC Florian

REMERCIEMENTS :

Remerciements de Élise AGUERA

À notre directeur de thèse, Monsieur le Professeur Jordan BIREBENT, vous nous avez fait l'honneur de diriger notre thèse. Je vous remercie pour la confiance que vous nous avez accordée durant ce travail, de votre disponibilité et de la bienveillance que vous avez eu à notre égard. Vos conseils éclairés, notamment au niveau méthodologique, nous ont permis de surmonter bien des obstacles.

À Madame la Présidente du Jury, Madame la Professeur Julie DUPOUY, vous nous faites l'honneur de présider ce jury et de juger notre travail. Veuillez trouver ici l'expression de ma gratitude.

À Monsieur le Professeur Jean-Christophe POUTRAIN, votre dévotion auprès des internes et de la faculté force l'admiration. Nous vous remercions de juger ce travail.

À Monsieur Jean-Marie BARBICHE, Responsable du Département Appui à la Recherche et Science Ouverte de l'Université Toulouse III, je vous remercie pour votre disponibilité et votre pédagogie. Votre assistance lors de nos recherches bibliographiques nous a été précieuse, tout comme vos conseils de recours à vos collègues de l'Université de Paris Cité. Je souhaite également accorder mes remerciements à Anne-Morgan DEVRIENDT et Colin SIDRE de la BIU Santé Pharmacie et Médecine de l'Université Paris Cité pour la réalisation de la recherche bibliographique sur la base de données Embase.

À Muriel FERRÉ, un grand merci pour nous avoir aidé dans la mise en page de ce travail.

Je tiens à exprimer toute ma gratitude à mes maitres de stage, qui m'ont accueillie et encadrée lors de mes trois années d'internat. Merci au Dr Patrick SALGUES pour ta gentillesse lors de ce mémorable premier jour, et tous ceux qui ont suivi, sans compter les incroyables discussions que nous pouvions avoir lors des consultations. Que ce soit sur la mécanique, la gastronomie, la musique ou la région, tu étais un véritable puit d'informations. Je remercie également le Dr Hervé GUILLEUX qui de par son vaste éventail de compétences et son intérêt dans les sciences m'a appris que la médecine générale n'avait pour limites que celles que l'on se fixait à nous-mêmes.

Je remercie le Dr Fanny MORIN pour son accueil et son encadrement dans le beau département de l'Aveyron. Merci également au Dr Delphine MAYEUR pour ta bienveillance et ta patience lors des gestes de gynécologie. Tu as fait de l'apprentissage de cette discipline un bonheur, ce qui n'était pas gagné. Merci également à Charline VALROFF, et Pauline WOESTELANDT de m'avoir intégrée à la maison de santé de Figeac lors de ce semestre. Je souhaiterais adresser une pensée aux membres du service de Pédiatrie du Centre Hospitalier de Rodez pour avoir partagé avec tant de patience leurs connaissances.

Un grand merci à l'équipe du SSR de Vic en Bigorre, auprès de qui j'ai appris énormément sur le plan thérapeutique, mais également humain. Je voudrais exprimer ma reconnaissance envers le Dr Yannick GASNIER dont les connaissances et compétences en gériatrie forcent l'admiration. Merci au Dr Alexandre CHARNAVEL pour ton enthousiasme et ta soif continue d'apprendre, qui ont été un véritable exemple. Je te dois mes progrès en optimisation thérapeutique, et pour cela, je te suis infiniment reconnaissante. Merci également aux urgentistes de Tarbes, et plus particulièrement le Dr Jules CLOUZEAU pour son efficacité et sa disponibilité envers les internes.

Je souhaite remercier chaleureusement mes maitres de stage de dernière année d'internat, Dr ROUSSEL, Dr VALOBRA, Dr BORDES, Dr REMONDET, Dr BLOCH-PLOUVIER, Dr SAVIGNAC et Dr MALLET-RAOULT. Vous m'avez beaucoup apporté humainement et scientifiquement, et je suis fière de pouvoir exercer en tant que docteur grâce à votre encadrement bienveillant. Vous avez guidé mes pas lors de mes premières consultations en autonomie et je vous en resterai reconnaissante.

Pour Nicolas, merci de ta gentillesse et de ton soutien. Je n'oublierai pas les délicieux repas au prieuré où nous pouvions parler de tout et qui, quelque part, furent le point de départ de cette thèse.

Pour Marie, ton accueil et ta bienveillance m'ont permis de prendre confiance en moi, revenir te remplacer a été un plaisir.

À Fabienne, tu as été le rocher auquel s'est raccrochée la collocation de Figeac pendant ce semestre d'été 2023. Te rencontrer a été une chance exceptionnelle, et échanger avec toi des prises en charge un enrichissement sans mesure. Le cabinet REMONDET est devenu pour moi une deuxième maison, avec Louise et Chrystelle et c'est avec plaisir que je reviens pour y remplacer. Je n'oublierai pas non plus nos consultations de gynécologie au CH de Figeac, un véritable bonheur.

Une pensée pour Nathalie, que je n'ai connu que trois mois. Tu avais tant à nous apprendre, mais la vie en a voulu autrement. Je me rappellerai de nos conversations sur le balcon de la maison de santé où tu nous racontais tes incroyables anecdotes.

Pour Jérôme, toi qui m'as appris tant de choses en médecine du sport. Ces fameuses matinées au CH de Figeac ont été une expérience très formatrice et je suis heureuse de pouvoir prétendre en avoir retenu quelques bases que j'appliquerai à ma pratique.

Au Dr Nathalie GIACCO-DEMOULINS, j'avais sept ans lors de notre première rencontre. Sous un fond de suivi de consultations de dermatologie, un véritable lien s'est créé. Ta passion pour la médecine et ton niveau d'expertise ont toujours forcé mon admiration. Tu m'as toujours apporté de précieux conseils et encouragements lors de mes années universitaires, notamment face à la première année commune de santé et face à mes doutes quant au lieu de réalisation de mon internat. Apprendre à tes côtés est un véritable plaisir, que je renouvelle avec enthousiasme. Un grand merci pour tout.

Je voudrais également remercier mes amis, qui m'ont tant appris et apporté.

Aux filles de la collocation des Flou'zzz, merci d'avoir été présentes pendant ces premières années d'études. J'ai appris énormément de la vie à vos côtés, la vie en communauté, l'entretien d'une maison, les courses... Cette entraide nous a permis d'atteindre nos buts ensemble, vous aurez toujours une place dans mon cœur malgré les chemins différents pris par chacune.

Petite dédicace spéciale à Alexia, nous avons partagé tellement de choses ! Tout ça a commencé par notre rituel du matin à la cafétéria (Y... Quinze !). Je n'en reviens pas du chemin parcouru ! Merci d'avoir été là pendant toutes les périodes d'exams, lorsque je perdais confiance en moi. Tu avais toujours les mots justes pour m'encourager. Merci aussi pour ton brin de folie, qui associé au mien m'a permis d'avoir le courage de continuer une pratique sportive au sein du MHC.

À Lianne, tu as été ma première amie à la UAX. Nos conversations en anglo-espagnol, complètement loufoques me font encore sourire. Malgré des parcours professionnels différents et la distance, tu es une véritable amie. Je te souhaite beaucoup de bonheur !

Au groupe de MirAsturias, Alvaro, Cintia, Maria C., Claudia et Maria. Ces quelques mois de préparation pour le MIR furent une vraie folie ! Un pari incroyable que nous avons réussi ensemble. Je me souviens de toutes ces heures en soirée passées à la prépa, sans compter ces moments de correction avec Alvaro, notre professeur particulier !

À Cintia, je te félicite pour cette fin d'internat à Laredo. L'énergie, la gaité et le dynamisme dont tu fais preuve dans tous les aspects de ta vie sont une bouffée d'oxygène. A très bientôt pour nos aventures à cheval entre la France et l'Espagne.

Maria C. je te souhaite beaucoup de bonheur en Galice, pour ce nouveau poste dont j'espère bientôt avoir plus de nouvelles. Mais surtout, j'espère te revoir au prochain congrès de Médecine Générale.

Alvaro, tu es la bonté et la bienveillance même. Votre petite Alicia a bien de la chance d'avoir des parents aussi exceptionnels, et le Puerta un médecin en or. Profite bien de ton métier de Papa.

Aux membres de l'AIMG-MP (avec dédicace spéciale à Carole et Blandine), vous avez été, lors de ce début d'internat en plein Covid, un véritable noyau. Je me rappellerai longtemps de ces réunions d'assemblées générales que je suivais depuis mon lieu de stage de PN1 dans le Lot. Je vous souhaite tout le bonheur du monde. A bientôt pour la soutenance de Carole et Blandine !

À Emily et Angélique, que je remercie pour leur amitié et nos après-midis d'escalades, j'espère partager bien d'autres aventures avec vous.

À Jacque et Richard, sans qui mon parcours aurait été bien différent. Merci pour votre soutien, vos précieux conseils et vos marques d'amitiés constantes.

À Cécile, mon amie, comment résumer ce qu'on a vécu en quelques mots ? Cette histoire a comme début un partage de nos lieux de stages lors de notre première année d'internat. Ce chassé-croisé nous avait permis d'échanger par messages sans nous rencontrer, et déjà je sentais que tu étais une personne exceptionnelle. Et puis le hasard nous a réunies pour une cohabitation de deux semestres et quelle cohabitation ! Pendant un an, vivre et échanger ces expériences de stage avec toi a été un bonheur, malgré les difficultés de l'internat. Et lorsque je t'ai proposé cette thèse, tu m'as fait confiance sans la moindre hésitation et tu t'es plongée à fond dans ce projet ! Je me considère comme chanceuse de t'avoir rencontrée et de pouvoir t'appeler mon amie. Merci infiniment.

À mes parents, que j'aime tendrement. La langue Française ne comporte pas assez de mots pour que je vous exprime tout ce que je ressens à votre égard. Ce que je suis aujourd'hui, je vous le dois. Vous m'avez soutenue à travers toutes les épreuves de la vie, et lorsque j'ai voulu persévérer dans la voie de la médecine, vous avez cru en moi. Votre foi et votre amour inconditionnels me transportent. La reconnaissance et l'amour que j'éprouve à votre égard ne peuvent être mesurés.

À ma sœur Mathilde, mon associée de toujours dans toutes mes aventures. A toi qui m'as accompagnée au bout du monde comme sur les pistes de ski, ta force, ton dynamisme et ton énergie forcent mon admiration. J'ai hâte de pouvoir assister à mon tour au couronnement de tes efforts lors de ton mémoire.

À mon grand-père, Albert-Pierre dit « Pépé ». Tu as guidé toutes mes aspirations. Le violon, la musique, la médecine, tu as été ma source d'inspiration. J'espère que d'où tu es, je continuerais à faire ta fierté.

Remerciements de Cécile FERRÉ

À notre Directeur de thèse, le Pr Jordan BIREBENT, merci d'avoir accepté de nous accompagner dans ce travail d'analyse systématique de la littérature. Toujours avec bienveillance, votre soutien, en particulier sur le plan méthodologique, a été d'une aide précieuse pour nous recentrer sur les étapes essentielles de l'élaboration de ce manuscrit. Merci pour votre expérience, votre gentillesse et pour l'honneur que vous nous avez fait.

Au Professeur Julie DUPOUY, un grand merci d'avoir accepté d'évaluer ce travail et de présider mon jury. Vous avez toute ma gratitude, y compris pour votre implication importante dans la Recherche en médecine générale et dans l'Enseignement aux internes.

Au Professeur Jean-Christophe POUTRAIN, merci de nous faire l'honneur de juger ce travail et d'avoir été disponible pour la soutenance. Je tiens également à vous remercier pour vos cours, particulièrement enrichissants, dispensés durant mon internat.

Au Dr Nicolas ROUSSEL, merci d'avoir partagé avec Élise l'idée de travailler sur l'IA, une suggestion qui m'a beaucoup enthousiasmée. Je vous remercie pour l'honneur de votre présence à la présentation de ce travail.

Mes remerciements chaleureux à Monsieur Jean-Marie BARBICHE, Responsable du Département Appui à la Recherche et Science Ouverte de l'Université Toulouse III, pour son aide précieuse au cours de nos recherches bibliographiques sur *PubMed*. Remerciements également à Anne-Morgane DEVRIENDT et Colin SIDRE de la BIU Santé Pharmacie et Médecine de l'Université Paris Cité pour leur travail de recherche sur *Embase*.

Je tiens à exprimer ma reconnaissance à mon ancienne équipe du laboratoire de l'INSERM/Purpan, qui m'a accompagnée lors de ma première thèse en biologie. Un merci tout particulier au Dr Daniel DUNIA, un chef patient et humain, qui m'a permis de réaliser des stages en médecine afin d'obtenir la passerelle. Sans votre soutien, je ne serais pas ici aujourd'hui. Merci également au Dr Marion SZELECHOWSKI, mon encadrante lors de mon master et de ma thèse, pour ses précieux conseils scientifiques et sa disponibilité. Un grand merci à "Maman" Anne, ingénieure exceptionnelle, qui m'a accompagnée dans tant d'expériences, au sens propre comme au figuré. Merci pour tout ce que j'ai appris auprès de vous.

Ma gratitude va également au Professeur Hugues CHAP, pour m'avoir fait l'honneur de me recommander pour mon entrée en médecine et pour avoir assisté à ma première soutenance de thèse.

Un immense merci à tous les professeurs de l'Université d'Orléans, de Toulouse Paul Sabatier puis de Toulouse Purpan et Rangueil, pour leur transmission de connaissances et leur capacité à inspirer tant d'étudiants, dont moi-même.

Je tiens à exprimer ma reconnaissance à tous mes maîtres de stage durant ces trois années d'internat. Merci au Dr Hervé GUILLEUX, pour tout ce que tu m'as appris scientifiquement et en gestes pratiques, ainsi que pour nos discussions philosophiques enrichissantes lors de nos trajets « interminables » mais au combien précieux au fin fond du Lot en VAD. Un merci particulier au Dr Patrick SALGUES, "un médecin à l'ancienne", tellement humain, avec un flair incroyable, drôle et accueillant. J'ai été ravie de passer deux semestres à tes côtés.

À toute l'équipe des urgentistes de CASTRES, merci infiniment pour votre savoir et pour cette expérience inoubliable. Un énorme merci au Dr Pierre HARMAND, pour son efficacité légendaire, sa pédagogie et sa compréhension vis-à-vis de nous, jeunes internes. Merci également aux Dr Jean-Louis MONTSERRET et Dr Clément CANIFFI pour leur partage de connaissances, leur calme en toutes circonstances, et bien sûr, pour les fous rires partagés. Je remercie également le Dr BOULARAN pour son accueil dans son service et ses conseils avisés, ainsi que les Dr GUILLON, Dr GOSSELIN, Dr LE COZ, Dr SCHNEIDER, Dr RICARD et Dr ROUX, pour la qualité de leur encadrement et leur bienveillance.

Un grand merci à la Dr Solène LESPINASSE pour ce semestre dans la magnifique région de l'Aveyron, pour tout ce que j'ai acquis et pour ces moments de partage, même dans les périodes difficiles. Merci également au Dr Marielle PUECH pour ta gentillesse et la qualité de tes enseignements.

Je tiens également à remercier l'équipe de Médecine Interne du CH de Tarbes, qui m'a beaucoup appris, ainsi que les Urgentistes de Tarbes. Malgré un semestre difficile, ce fut une expérience très formatrice. Un immense merci au Dr Yannick GASNIER, pour son expertise en gériatrie et sa profonde gentillesse. Merci au Dr Alexandre CHARNAVEL, pour ton accueil, nos fous rires et nos journées de travail dans une ambiance sereine. J'espère que tu profites pleinement de tes montagnes alpines. À bientôt !

Pour ma dernière année d'internat, je remercie chaleureusement mes MSU Iotois, Dr RUDZINSKI, Dr ALRIC et, de nouveau, Dr Patrick SALGUES pour tout ce qu'ils m'ont apporté. Grâce à vous, j'ai gagné en autonomie et grandi en tant que jeune docteur. Je suis ravie d'avoir adopté le Lot et je sais que nous nous recroiserons bientôt. Ma gratitude va également à mes trois derniers MSU : Dr Nathalie BOUSSIER, Dr Denis PORTE à Montricoux et Dr Cécile ERHEL à Caylus dans le Tarn-et-Garonne. Merci pour ce dernier semestre si serein, rempli de bons conseils et de bienveillance. J'aime tant venir vous remplacer, on se sent si bien chez vous ! Je vous souhaite plein de joies et de bonheur : une "presque pré-retraite" bien méritée pour Denis, de magnifiques découvertes avec Mycroft pour Nathalie et de belles émotions avec Doudou, Dieyna et Mahé pour Cécile. Un merci tout particulier aux secrétaires (sans qui rien ne tournerait aussi bien et qui m'ont été si souvent d'une grande aide dans la prise en charge des patients) Morgane, Juliette, Jocelyne et Patricia, ainsi qu'au magnifique duo de Caylus, Estelle et Gérard.

Enfin, je remercie bien entendu les Dr Thierry MEURIE et Dr Nathalie CHAOUI, qui m'ont fait découvrir les soins palliatifs et les consultations douleur tout au long de ma dernière année d'internat lors de cette fameuse journée « annexe ». Merci de m'avoir tant appris et fait confiance. Ce fut une véritable vocation pour moi et je suis honorée de poursuivre cette aventure à vos côtés. Un immense merci à toute l'équipe "Soins Pall" : Bertrand, les 2 magnifiques IDE Aurélie et Aurélie, Émilie, Angélique et « notre » super secrétaire Sophie.

Que serait la vie sans les amis ? Par où commencer ? Comment tout vous dire et vous remercier ?

Pierre et Manu, mes deux amis fidèles du lycée, je suis tellement heureuse d'avoir croisé votre route, et d'être toujours à vos côtés 20 ans plus tard, malgré la distance. À très vite pour notre séjour à l'orthocentre !

À la fine troupe d'Orléans, Camille, Insaf, Naïm, Jeffrey, vous avez rendu la Licence tellement joyeuse. Quentin, mon "p'tit loup", inspecteur des fraudes et jeune papa d'un adorable petit garçon, Théodore, je te souhaite plein d'amour dans cette nouvelle aventure.

À la belle KhaKha, qui m'a supportée durant le master 1 et bien après, je te souhaite tout le bonheur dans ton nouveau rôle de maman. Bienvenue à Ali, qui a une maman extraordinaire !

Un énorme bisou aux filles de la thèse de bio, mon groupe des "bitches", de véritables amies : Céline, merci pour ces soirées pleines de joie et de jeux. Sophie (dont j'entends le rire communicatif à des milliers de kilomètres), je suis ravie de te revoir arriver en terre française, après avoir bravé le froid suédois. Beaucoup de bonheur pour ton nouveau poste de grande chercheuse amplement mérité ! Anne, ma "baby", je suis tellement fière de la femme si intelligente que tu es devenue. Tu as quitté la recherche pour l'informatique et je suis heureuse de te voir si épanouie. À très bientôt pour de nouvelles aventures. Et Mimi, ma petite boule d'amour rencontrée en M1, Lucie est un prénom qui te va si bien. Tu es une véritable source de lumière dans ce monde brutal. Merci d'être dans ma vie, et j'embrasse aussi Arnaud et Joseph. À très vite à Cambridge !

À toi ma belle Raïssa, également rencontrée lors de ma première thèse sur le plateau « imagerie », je te souhaite plein de bonheur pour cette nouvelle aventure un peu dingue de « prof d'SVT » à Mayotte !

À Daniella, Nanis, Vava et Marion, de superbes filles rencontrées aux Urgences de Castres. Je suis heureuse d'avoir partagé cette expérience avec vous. Malgré l'adversité nous sommes restées soudées. Bravo pour votre thèse, je vous souhaite tout le meilleur pour l'avenir.

À toi, Emily, douce et discrète, rencontrée également aux Urgences, sache que tu seras toujours la bienvenue chez moi !

Élise, mon amie et co-interne. Nous avons un peu échangé sur les réseaux sociaux avant de « se rencontrer en vrai » car nous avons effectué notre première année dans les mêmes stages mais « en alternance ». Je t'appréciais déjà beaucoup, puis vint La Rencontre, le destin qui a fait que nous avons pu partager les mêmes maisons d'interne pendant 1 an. Ce fut un véritable coup de cœur. Cela a été si simple de vivre ensemble (et on ne peut pas dire ça de tout le monde !). J'aime tant râler avec toi sur tout et refaire le monde ! Merci d'avoir proposé que je te rejoigne dans cette aventure un peu folle qu'est la thèse. J'ai commencé avec une collègue et je termine avec une véritable amie.

Enfin, un immense merci à ma famille, qui m'a soutenue sans relâche dans cette "folie" d'une deuxième thèse et d'un léger « rajout » de nombre d'années d'études. Il n'y a pas assez de place ici pour exprimer tout ce que je ressens. Si je suis ce que je suis aujourd'hui, c'est grâce à vous.

Je dédie cette thèse à Mémé Lucette, Pépé Robert et Tonton Nono. Merci d'avoir existé.

À Léa et Laura, mes nièces adorées, devenues adolescentes et bientôt femmes. J'adore passer du temps avec vous et j'espère être encore présente pour tous les moments forts de votre vie.

Maman, merci d'être là, pour ton aide, ton soutien infaillible et pour tous ces bons petits plats qui m'attendaient les soirs de dur labeur. Merci pour ta joie de vivre et ta façon unique de toujours croire en moi. Et pour toutes ces balades philosophiques (avec feu notre Crapiche d'amour, puis ses deux merveilles Nini et Gniouf, sans oublier notre Momo et bientôt les deux nouveaux venus !)

Mumu, cette thèse est aussi la tienne, encore une fois. Merci pour tes heures incalculables de mise en page, de corrections et de reformulations. Tout est déjà dit mais encore merci pour toutes ces années de suivi de devoirs depuis le collège, à me soutenir quand je suis « au bord du gouffre ». Sans toi, je ne serais jamais devenue "docteur-docteur".

Papa, mon Héros, ma Force, merci pour ton soutien indéfectible. Voir la fierté dans tes yeux est ce qu'il y a de plus beau. Et bien sûr, merci aussi à Mélo, la "Truffe", pour nous forcer à marcher des heures en forêt !

Enfin, Christian ("CriCri"), mon Amour, merci pour ton soutien sans faille durant ces 15 ans dépassés. Merci d'avoir supporté mes peurs et mes angoisses sans jamais te plaindre. C'est aussi grâce à toi que j'en suis là aujourd'hui.

Je vous aime.

TABLE DES MATIÈRES

LISTE DES ABRÉVIATIONS :	8
LISTE DES ILLUSTRATIONS :	9
RÉSUMÉ	10
SUMMARY	11
I - INTRODUCTION :	12
I-1 DÉMOGRAPHIE :	12
I-2 CANCERS CUTANÉS :	14
I-3 PRINCIPE DE L'INTELLIGENCE ARTIFICIELLE DANS L'ANALYSE DES LÉSIONS CUTANÉES PIGMENTÉES :	20
II - MATERIEL ET MÉTHODE :	33
II-1 TYPE D'ÉTUDE :	33
II-2 SOURCES D'INFORMATIONS ET STRATÉGIE DE RECHERCHE :	33
II-3 CRITÈRES DE SÉLECTION ET D'EXCLUSION :	33
II-4 SÉLECTION DES ÉTUDES :	34
II-5 ÉVALUATION DES ARTICLES SÉLECTIONNÉS :	34
III - RÉSULTATS :	36
III - 1 SÉLECTION DES ARTICLES :	36
III - 2 ARTICLES INCLUS ET RÉSULTATS PRINCIPAUX :	38
III - 3 ANALYSE DES RÉSULTATS	50
a) Analyse des lésions cutanées grâce à l'IA intégrée dans des applications pour Smartphones :	53
b) Évaluation de la performance d'une forme d'AI dans l'analyse de lésions cutanées :	54
2-a) IA complexe :	55
2-b) IA et analyse d'images séquentielles :	58
2-c) IA et aide à la décision pour la prise en charge :	59
c) Performance d'une IA versus médecins (dermatologues et/ou médecins généralistes et/ou autres intervenants en soins primaires)	61
d) Évaluation de la coopération entre l'IA et les médecins (dermatologues, médecins généralistes ou autres)	62
IV - DISCUSSION	66
V - CONCLUSION	75
VI - RÉFÉRENCES :	76
VII - ANNEXES	83
VII- 1 ANNEXE 1 : Équations de recherche sur les trois bases de données utilisées	83
VII-2 ANNEXE 2 : Liste des articles exclus	91
VII- 3 ANNEXE 3 : Évaluation de la qualité et de la transparence des articles selon les échelles STROBE, TREND et CONSORT	93
VII-4 ANNEXE 4 : Évaluation de la méthodologie des articles selon les échelles QUADAS-2, MINORS et Jadad	101
VII-5 ANNEXE 5 : Évaluation de la transparence et de la rigueur du rapport de notre revue systématique de la littérature ainsi que de sa qualité méthodologie grâce aux échelles PRISMA et AMSTAR	104

LISTE DES ABRÉVIATIONS :

AA : Apprentissage Automatique

ABCDE : Asymétrie, Bord, Couleur, Diamètre, Évolution (Méthode de classification des lésions cutanées)

AKIEC : Actinic Keratoses and Intraepithelial Carcinoma (Kératoses Actiniques et Carcinome Intra-Épithélial)

AMSTAR : A Measurement Tool to Assess Systematic Reviews (Outil d'évaluation des revues systématiques)

ANN : Artificial Neural Network (Réseau de Neurones Artificiels)

AP : Apprentissage Profond

AUROC : Area Under the Receiver Operating Characteristic Curve (Aire Sous la Courbe ROC)

BCC : Basal Cell Carcinoma (Carcinome Basocellulaire)

BKL : Benign Keratosis Lesions (Lésions de Kératose Bénigne)

CAD : Computer-Aided Diagnosis (Diagnostic Assisté par Ordinateur)

CNN : Convolutional Neural Network (Réseau de Neurones Convolutif)

CONSORT : Consolidated Standards of Reporting Trials (Normes Consolidées de Présentation des Essais)

DF : Dermatofibrome

GAN : Generative Adversarial Network (Réseau de Neurones Antagonistes Génératifs)

HAM : Human Against Machine (Base de données d'images de lésions cutanées)

HAS : Haute Autorité de Santé

IA : Intelligence Artificielle

INCa : Institut National du Cancer

ISIC : International Skin Imaging Collaboration (Collaboration Internationale d'Imagerie Cutanée)

KNN : K-nearest neighbors algorithm (Algorithme des Plus Proches Voisins)

MEL : Mélanome

ML : Machine Learning (Apprentissage Automatique)

MINORS : Methodological Index for Non-Randomized Studies (Indice Méthodologique pour les Études Non Randomisées)

NMSC : Non-Melanoma Skin Cancer (Cancer de la Peau Non Mélanome)

NV : Naevus

OMS : Organisation Mondiale de la Santé

PRISMA : Preferred Reporting Items for Systematic Reviews and Meta-Analyses (Éléments Préférés pour les Rapports des Revues Systématiques et Méta-Analyses)

QUADAS-2 : Quality Assessment of Diagnostic Accuracy Studies (Évaluation de la Qualité des Études de Précision Diagnostique)

RNN : Recurrent Neural Networks (Réseaux de Neurones Récurrents)

RvNN : Recursive Neural Networks (Réseaux de Neurones Récurifs)

SFD : Société Française de Dermatologie

SDD : Sequential Digital Dermoscopy (Dermoscopie Numérique Séquentielle)

SDT : Score Dermatoscopique Total

SVM : Support Vector Machine (Machine à Vecteurs de Support)

STROBE : Strengthening the Reporting of Observational Studies in Epidemiology (Renforcement du Rapport des Études Observationnelles en Épidémiologie)

TREND : Transparent Reporting of Evaluations with Nonrandomized Designs (Rapport Transparent des Évaluations avec des Designs Non Randomisés)

VASC : Lésions Vasculaires

LISTE DES ILLUSTRATIONS :

SOMMAIRE DES TABLEAUX :

Tableau 1 : Score dermatoscopique total calculé à partir des critères ABCD

Tableau 2 : Critères dermoscopiques de la liste en 7 points

Tableau 3 : Algorithme de Menzies

Tableau 4 : Caractéristiques et évaluation des études sélectionnées

Tableau 5 : Classement des articles sélectionnés par thème

SOMMAIRE DES FIGURES :

Figure 1 : Carte de la répartition géographique des dermatologues de moins de 40 ans (à gauche) et de plus de 60 ans (à droite) sur le territoire Français

Figure 2 : Pyramide des âges des spécialistes en dermatologie et vénéréologie

Figure 3 : Mélanome en vue macroscopique puis dermoscopique

Figure 4 : Carcinome basocellulaire nodulaire

Figure 5 : Stratégie diagnostique en deux étapes des lésions pigmentées

Figure 6 : Algorithme utilisé pour différencier les lésions mélanocytaires des lésions non mélanocytaires

Figure 7 : Images dermoscopiques de naevus

Figure 8 : Images dermoscopiques de kératoses séborrhéiques

Figure 9 : Images dermoscopiques de carcinomes basocellulaires

Figure 10 : Images dermoscopiques d'angiomes et d'un naevus bleu

Figure 11 : Schématisation des différents patrons globaux selon l'analyse des patrons

Figure 12 : Neurone artificiel simulant le comportement d'un neurone biologique

Figure 13 : Principe d'une machine à vecteur de support

Figure 14 : Organigramme de la classification de trois maladies de la peau par SVM

Figure 15 : La structure du système CAD (Computer-Aided Design ou Conception Assistée par Ordinateur)

Figure 16 : Modèle d'arbre de décision

Figure 17 : Schéma de l'architecture du modèle Bayésien DenseNet-169

Figure 18 : Représentation architecturale globale du modèle de détection de mélanome

Figure 19 : Segmentation de lésions grâce au KNN

Figure 20 : Extraction des informations de l'image grâce à un enchaînement de filtres (et de simplifications) puis prédiction de la classe de l'image

Figure 21 : Diagramme de flux de sélection des articles.

Figure 22 : Le CanTest Framework. D'après Jones et al

RÉSUMÉ

Contexte : Les cancers cutanés mélanocytaires et non mélanocytaires sont en forte augmentation dans le monde, en particulier en Australie. La France n'est pas épargnée par ce phénomène, l'incidence de ces cancers ayant triplé entre 1990 et 2023. 17 922 nouveaux cas de cancers mélanocytaires avaient été recensés en 2023 sur le territoire métropolitain. Ce phénomène est attribué à l'allongement de la durée de vie et à des expositions solaires inappropriées. L'accès aux consultations en dermatologie devient de plus en plus difficile avec un délai moyen de 64 jours, et la densité de dermatologues diminue constamment. Les médecins généralistes sont souvent sollicités pour diagnostiquer des lésions cutanées, bien qu'ils se sentent insuffisamment formés dans ce domaine. Cette situation a encouragé la recherche de solutions alternatives, telles que l'intelligence artificielle (IA), pour améliorer le diagnostic et la prise en charge des cancers cutanés.

Objectif : L'objectif principal de cette étude est d'évaluer si l'IA peut être utilisée pour aider les médecins généralistes à diagnostiquer et prendre en charge précocement les cancers de la peau, en particulier en soins primaires. L'étude vise aussi à recenser les différentes méthodes d'IA, évaluer leur efficacité, coût et accessibilité en comparaison avec les interventions des spécialistes dermatologues.

Méthode : Nous avons réalisé une revue systématique de la littérature, selon les recommandations internationales de l'échelle Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA). Les bases de données *Medline*, *Embase* et *Web of Science* ont été explorées grâce à un algorithme d'une étude du Lancet réalisée en 2021, pour identifier les études publiées entre août 2021 et août 2023. Les critères d'inclusion ont été choisis pour sélectionner des études portant sur l'utilisation d'algorithmes d'IA appliqués au diagnostic du cancer de la peau, en particulier en soins primaires. Une élimination des articles ne correspondant pas au sujet d'étude a été réalisée par étapes, tout d'abord par lecture des titres, puis des résumés et enfin des textes complets à deux chercheurs. La qualité des rapports a été évaluée à l'aide des échelles STROBE, TREND et CONSORT. L'évaluation de la qualité méthodologique des études a été effectuée avec les grilles QUADAS-2, MINORS et Jadad.

Résultats : Sur 2 046 articles identifiés, 25 études ont été incluses dans la revue systématique. Parmi elles, 25 % portaient sur les soins primaires et rapportaient des résultats discordants sur la précision diagnostique de l'IA par rapport à celle des médecins généralistes dans un environnement clinique. Les études portant sur la coopération de médecins (généralistes ou dermatologues) avec l'IA démontraient une amélioration de l'exactitude et de la sensibilité de diagnostic. Certains algorithmes, comme les réseaux de neurones convolutifs (CNN), ont montré des performances diagnostiques similaires, voire supérieures à celles des dermatologues dans des environnements contrôlés. Les approches d'IA multimodales combinant images et métadonnées (âge, sexe...), les algorithmes d'apprentissage profond (entraînés sur des images standardisées) ainsi que l'apprentissage par renforcement (ajustant leurs décisions en fonction des retours d'expérience), ont significativement amélioré la précision diagnostique, notamment pour les mélanomes.

Discussion : Les résultats indiquent que l'IA pourrait jouer un rôle important dans l'amélioration de la détection précoce des cancers cutanés en soins primaires, réduisant ainsi le recours aux spécialistes dans certains cas. Néanmoins, les méthodes d'IA existantes à ce jour nécessitent encore des validations cliniques plus robustes et leur généralisation aux contextes réels pose des défis, notamment en ce qui concerne les populations sous-représentées dans les ensembles de données d'entraînement. Des questions restent ouvertes sur la reproductibilité des résultats en dehors des environnements de tests théoriques, ainsi que sur l'acceptabilité par les praticiens et les patients.

SUMMARY

Background: Melanocytic and non-melanocytic skin cancers are increasing worldwide, particularly in Australia. France is not spared from this phenomenon, with the incidence of these cancers tripling between 1990 and 2023. In 2023, 17,922 new cases of melanocytic cancers were recorded in metropolitan France. This phenomenon is attributed to longer life expectancy and inappropriate sun exposure. Access to dermatology consultations is becoming increasingly difficult, with an average waiting time of 64 days, and the density of dermatologists is constantly decreasing. General practitioners are often called upon to diagnose skin lesions, although many feel insufficiently trained in this field. This situation has encouraged the search for alternative solutions, such as artificial intelligence (AI), to improve the diagnosis and management of skin cancers.

Objective: The primary objective of this study is to assess whether AI can be used to assist general practitioners in diagnosing and managing skin cancers early, particularly in primary care. The study also aims to review the different AI methods, evaluate their effectiveness, cost, and accessibility compared to the interventions of dermatology specialists.

Method: We conducted a systematic literature review following the international guidelines of the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) scale. The Medline, Embase, and Web of Science databases were explored using an algorithm from a Lancet study conducted in 2021, to identify studies published between August 2021 and August 2023. The inclusion criteria were selected to focus on studies related to the use of AI algorithms applied to skin cancer diagnosis, particularly in primary care. Irrelevant articles to the study were excluded in stages, first by reading the titles, then the abstracts, and finally the full texts by two researchers. The quality of the reports was assessed using the STROBE, TREND, and CONSORT scales. The methodological quality of the studies was evaluated using the QUADAS-2, MINORS, and Jadad checklists.

Results: Of the 2,046 articles identified, 25 studies were included in the systematic review. Among these, 25% focused on primary care and reported conflicting results on the diagnostic accuracy of AI compared to that of general practitioners in a clinical setting. Studies examining the cooperation between doctors (general practitioners or dermatologists) and AI demonstrated an improvement in diagnostic accuracy and sensitivity. Some algorithms, such as convolutional neural networks (CNNs), showed diagnostic performance similar to or even better than dermatologists in controlled environments. Multimodal AI approaches combining images and metadata (age, sex, etc.), deep learning algorithms (trained on standardized images), and reinforcement learning (adjusting decisions based on feedback) significantly improved diagnostic accuracy, especially for melanomas.

Discussion: The results suggest that AI could play an important role in improving the early detection of skin cancers in primary care, thus reducing the need for specialists in some cases. However, the current AI methods still require more robust clinical validations, and their generalization to real-world contexts presents challenges, particularly concerning underrepresented populations in training datasets. Questions remain about the reproducibility of results outside of theoretical test environments, as well as the acceptability of AI to practitioners and patients.

I - INTRODUCTION :

I-1 DÉMOGRAPHIE :

Dans le monde entier, les cancers de la peau, mélanocytaires et non mélanocytaires, sont en forte expansion. C'est particulièrement vrai et important en Australie, où l'incidence de ces lésions a atteint un pic inégalable. La France n'y échappe pas, on a ainsi pu constater une incidence multipliée par trois entre 1990 et 2023. Rien que pour l'année 2023, 17 922 nouveaux cas de cancers mélanocytaires ont été recensés en France métropolitaine (1).

Cet accroissement du nombre de cancers cutanés est dû à une augmentation de la longévité mais également à des expositions solaires inappropriées au cours de la vie de chaque individu. Les demandes d'expertise médicale, aux fins de diagnostics, augmentent donc perpétuellement en parallèle. Or, les patients ont de plus en plus de difficultés à obtenir un rendez-vous en dermatologie (1,2).

Selon l'observatrice d'accès aux soins, en 2017, le délai moyen de rendez-vous chez un spécialiste en dermatologie est de 64 jours, soit 23 jours de plus qu'en 2012. Ces chiffres en font une des spécialités les plus difficiles d'accès, les délais moyens en cardiologie et rhumatologie sont estimés, respectivement, à 51 et 53 jours. Notons que l'accès moyen à une consultation en médecine générale est évalué à 8 jours (2).

La faible proportion de dermatologues sur le territoire explique cette accessibilité restreinte. En 2023, nous comptons une proportion de 3,4 dermatologues pour 100 000 habitants. A ceci, s'ajoute un intérêt croissant pour des actes relevant de l'esthétique, ce qui réduit les consultations pour des motifs médicaux. (3,4)

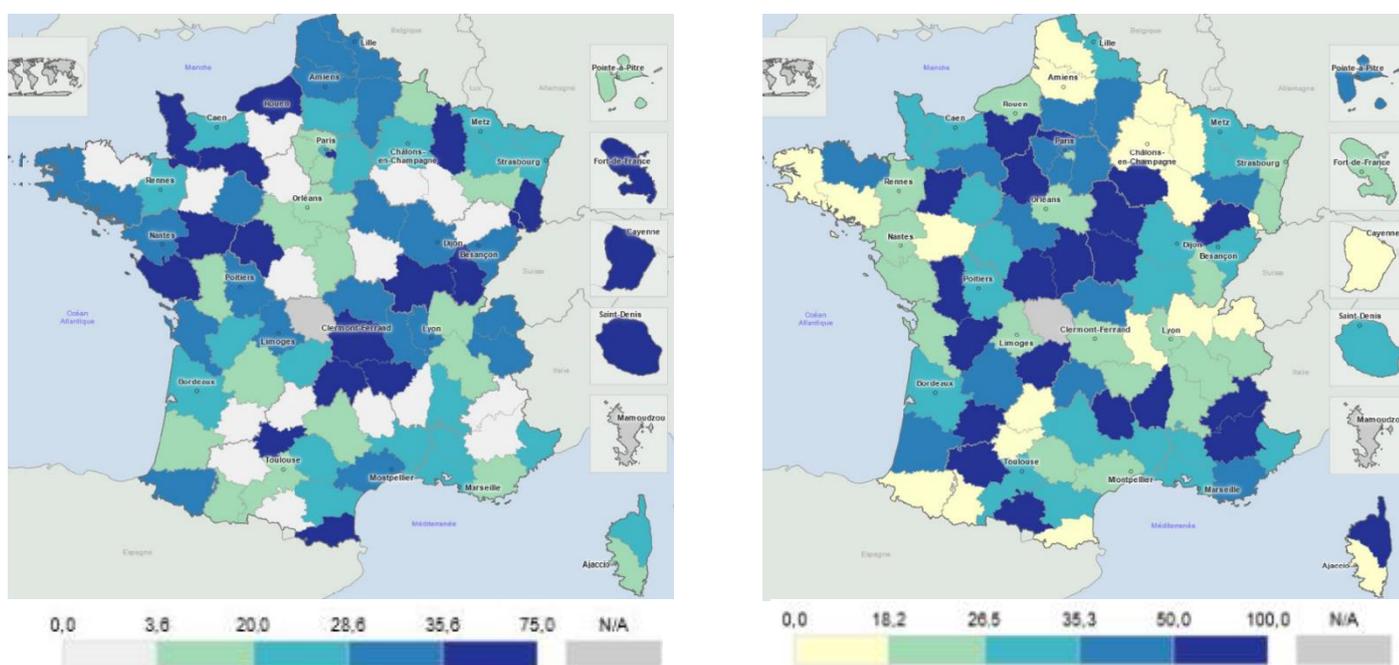


Figure 1 : Carte de la répartition géographique des dermatologues de moins de 40 ans (à gauche) et de plus de 60 ans (à droite) sur le territoire Français, par Arnault F. (3). Nous pouvons observer la densité de dermatologues par département, ceux en bleu foncé étant les mieux pourvus et en gris ou jaune les moins dotés.

Cette population de spécialistes tend au vieillissement, comme nous pouvons l'observer sur la figure n°2. Plus de la moitié des effectifs des dermatologues actifs en 2023 est âgée de plus de 50 ans (3).

Toujours en 2023, le nombre de dermatologues actifs était de 2 981 selon l'ordre des médecins, contre 3 918 en 2018. Il est par ailleurs estimé que la diminution des effectifs de spécialistes en dermatologie devrait perdurer dans les 10 prochaines années, leur nombre étant prédit à 2717 en 2030 (2,3).

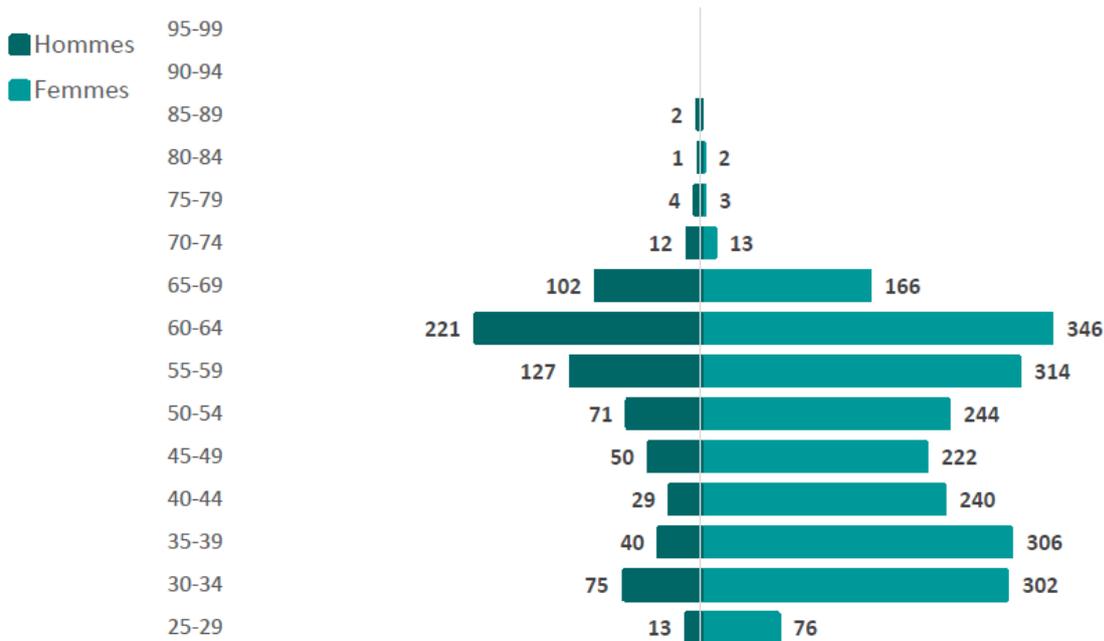


Figure 2 : Pyramide des âges des spécialistes en dermatologie et vénérologie par Arnault F. (3).

En conséquence de cette accessibilité difficile au médecin spécialisé en dermatologie, les patients se tournent de plus en plus souvent vers les médecins généralistes pour établir un diagnostic relatif à ces lésions cutanées. L'étude ECOGEN, réalisée entre décembre 2011 et avril 2021, indiquait que 5,7 % des consultations en médecine générale étaient consacrées à la dermatologie, 12,8 % d'entre-elles concernaient des tumeurs malignes et bénignes cutanées (5,6).

Lors de travaux de thèse précédents, il a été objectivé que les généralistes considèrent être insuffisamment formés en dermatologie et se sentent en difficulté devant les demandes de diagnostic sur lésions cutanées. Cette situation entraîne la recherche de solutions alternatives afin qu'ils puissent rendre un diagnostic fiable ou non délétère à leurs patients. De fait, la demande d'avis de téléconsultation augmente exponentiellement auprès des dermatologues. Ceux-ci peinent à répondre à la demande et/ou ne souhaitent pas exposer leur responsabilité sans examiner le patient (2).

I-2 CANCERS CUTANÉS :

Les cancers cutanés se divisent en deux catégories, les cancers mélanocytaires et les non mélanocytaires. Ces deux types de cancers se développent à partir de cellules présentes dans l'épiderme et possèdent des terrains et facteurs de risques communs (7,8).

Ils résultent d'une exposition au soleil ou aux ultraviolets artificiels, des antécédents de coup de soleil, d'un phototype dit clair (I/II), d'un nombre élevé de grains de beauté (syndrome du naevus atypique), des antécédents personnels ou familiaux de cancer cutané (1,7).

Les cancers mélanocytaires ou mélanomes, représentent 10 % des cancers cutanés. Particulièrement agressifs, ils sont connus pour leur capacité à métastaser rapidement. Ainsi, malgré leur faible incidence parmi les cancers cutanés, ils contribuent majoritairement à leur morbi-mortalité. Ils peuvent survenir de novo, c'est à dire sur une peau saine dans 70 à 80 % des cas, ou bien résulter de la transformation maligne d'un naevus préexistant (1).

La phase de croissance s'effectue tout d'abord superficiellement, au sein de l'épiderme, à partir d'une prolifération non contrôlée de cellules mélanocytaires, avant de progresser en profondeur. On distingue quatre formes anatomocliniques de mélanomes : mélanome à extension superficielle, mélanome nodulaire, mélanome de Dubreuil et mélanome acrolentigineux. Le mélanome à extension superficielle représente la forme la plus fréquemment observée, constituant 60 à 70 % des mélanomes diagnostiqués.

Le pronostic des patients a un lien direct avec l'épaisseur de la tumeur, les formes superficielles diagnostiquées précocement pouvant être traitées par une chirurgie d'exérèse seule. En revanche, dans le cas de métastases à distance, le pronostic à 5 ans est défavorable malgré les thérapies adjuvantes (8,9).

Le diagnostic se fait à partir d'une suspicion clinique, qui peut reposer sur plusieurs critères. Tout premièrement, la règle du « vilain petit canard », qui permet de considérer comme potentiellement malin tout naevus différant des autres. Les critères ABCDE sont également une aide au diagnostic et classent comme suspicion de mélanome toute lésion obtenant un score supérieur ou égal à 2. Ces critères sont l'asymétrie, des bords irréguliers, une couleur inhomogène, un diamètre supérieur à 6 mm et une évolution récente. La perception du patient a aussi son importance dans le dépistage des mélanomes, étant donné que près de 80 % de ceux-ci sont détectés par les patients. (7)



Figure 3 : Mélanome en vue macroscopique puis dermoscopique par Ring et al. (10)

90 % des cancers cutanés sont des cancers non mélanocytaires. On distingue les carcinomes basocellulaires ou épidermoïdes qui se développent à partir des kératinocytes.

Les carcinomes basocellulaires sont les plus fréquents et les moins agressifs. Ils se développent à partir des kératinocytes de la couche basale de l'épiderme, restant localisés, sans risque de métastase. Cliniquement les carcinomes basocellulaires peuvent prendre des morphologies variables, cependant ils s'initient principalement par une papule de petite taille, de bordure perlée avec des vaisseaux en surface avant de présenter une dépression ou une ulcération. On peut les classer en trois catégories principales selon leur aspect, nodulaire (60 %), superficiel et sclérodermiforme. Leur exérèse complète permet une guérison sans récurrence. Cependant, une prise en charge tardive, du fait de leur caractère expansif, peut mener à un préjudice fonctionnel et esthétique important (7).



Figure 4 : Carcinome basocellulaire nodulaire selon Collège de Dermatologie (7)

Les carcinomes épidermoïdes ou spinocellulaires se développent à partir des cellules squameuses de l'épithélium malpighien de la peau et sont plus rares que les carcinomes basocellulaires. Ils surviennent principalement de la transformation d'une lésion précancéreuse sur peau photo-exposée, bien qu'ils puissent également apparaître de novo. Les lésions pré-cancéreuses en question sont les kératoses actiniques, les plaies chroniques ou bien une cicatrice cutanée, notamment sur brûlure. Les carcinomes épidermoïdes se présentent comme des lésions d'apparence épaisse, squameuse et irrégulière évoluant vers une ulcération. Ils ont la capacité d'envahir les ganglions lymphatiques et de métastaser. Il est donc primordial de les diagnostiquer précocement (7).

Le diagnostic de cancer de la peau, qu'il soit mélanocytaire ou non mélanocytaire, repose principalement sur un examen visuel complet cutané. Afin d'améliorer la précision de cet examen, les dermatologues ont, au cours du siècle dernier, cherché à développer des outils de plus en plus performants à base de loupes grossissantes. Leon Goldman, un dermatologue américain, est connu pour avoir développé à partir de 1951 l'usage de manière systématique dans son examen des lésions pigmentaires cutanées de l'outil connu aujourd'hui comme dermatoscope (10).

Le dermatoscope est un outil simple, de petite taille facilement manipulable, composé d'une lentille grossissante et d'une source de lumière. La dermoscopie ou microscopie par épiluminescence est une

technique permettant un examen non invasif, in vivo, des structures de la peau jusqu'au derme réticulaire, invisible à l'œil nu. Elle facilite la différenciation entre lésion mélanocytaire et non mélanocytaire d'une lésion cutanée et permet d'en évaluer la nature bénigne ou maligne. Il a été démontré, dans la méta-analyse de Kittler et al., que l'examen dermoscopique, lorsque réalisé par un médecin formé à son usage, apportait une précision de diagnostic supérieure à l'examen visuel (11).

Depuis la généralisation de la dermatoscopie, plusieurs méthodes ont été développées afin d'aboutir à un diagnostic. En 2001, le comité de consensus international de dermatoscopie a établi des recommandations dans la stratégie de classification et de diagnostic des lésions pigmentées. Celle-ci se déroule en deux phases successives. La première consiste à différencier les lésions mélanocytaires des lésions non mélanocytaires. La seconde phase consiste, une fois le diagnostic de lésion mélanocytaire établi, à différencier si la lésion est bénigne, maligne ou suspecte (figure 5) (12).

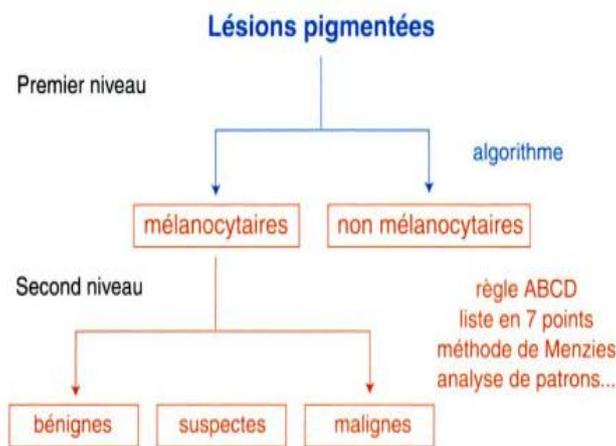


Figure 5 : Stratégie diagnostique en deux étapes des lésions pigmentées (12)

Pour différencier une lésion cutanée pigmentée en mélanocytaire ou non mélanocytaire, un algorithme en six étapes est proposé ci-dessous en figure 6.

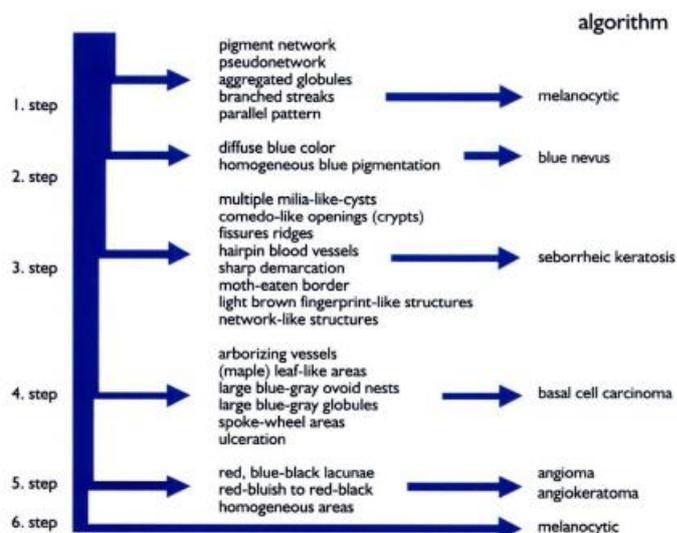


Figure 6 : Algorithme utilisé pour différencier les lésions mélanocytaires des lésions non mélanocytaires selon L.Tomas et R.P.Braun (12).

Première étape : identification d'un patron en faveur d'une lésion mélanocytaires ; globules, réseau pigmenté, stries ramifiées, pseudo-réseau pigmenté ou patron parallèle. Deuxième étape : identification en faveur d'un naevus bleu, pigmentation bleue homogène. Troisième étape : recherche d'une kératose séborrhéique ; identification de pseudo-comédons, pseudo-kystes cornés, fissures, aspect d'empreintes digitales, cérébriformes ou bordure mordillée. Quatrième étape : recherche de structures en faveur d'un carcinome basocellulaire ; identification de structures en « feuilles d'érable », roues dentées, grands-nids ovoïdes, points bleu-gris multiples, ulcérations ou vaisseaux sanguins ramifiés. Cinquième étape : identification de structures en faveur d'une lésion vasculaire ; globules rouges, rouge-bleus ou noirs. Sixième étape : au vu de l'absence de signes dermoscopiques spécifiques non-mélanocytaires repérés lors des précédentes étapes, la lésion est définie par défaut comme mélanocytaire pour la sécurité du patient.



Figure 7 : Images dermoscopiques de naevus selon L. Thomas et R.P. Braun (12). A gauche globules, au milieu réseau pigmenté et à droite naevus complexe avec stries ramifiées, globules et réseau.

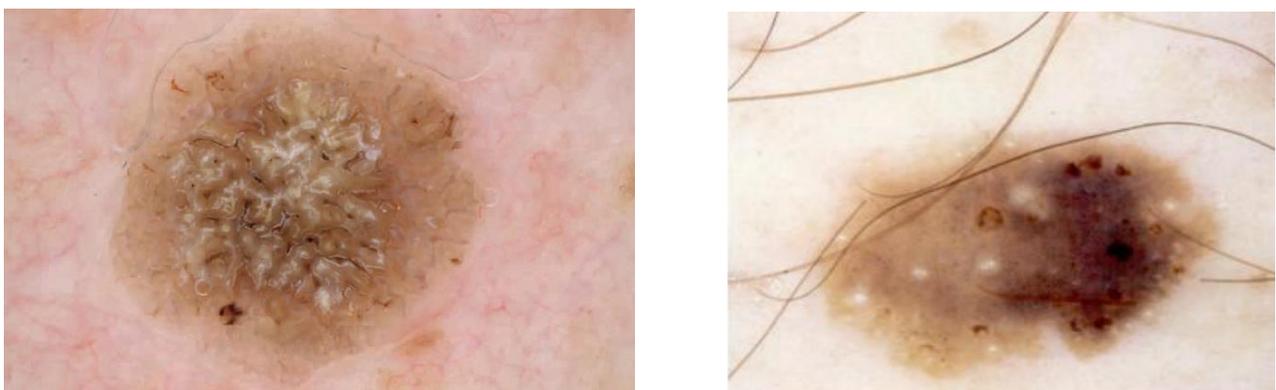


Figure 8 : Images dermoscopiques de kératoses séborrhéiques par L. Thomas et R.P Braun (12). Identification de structures cérébriformes dans l'image de gauche, de pseudo-kystes cornés et de pseudo-comédons dans celle de droite.



Figure 9 : Images dermoscopiques de carcinomes basocellulaires par L. Thomas et R.P. Braun (12) et DermoscopyAtlas (13). Structures de roue dentée, grands-nids ovoïdes, points bleu-gris multiples, vaisseaux arborescents.



Figure 10 : Images dermoscopiques d'angiomes et d'un naevus bleu par L. Thomas et R.P. Braun (12) et DermoscopyAtlas (13).

Parmi les algorithmes de décision de malignité des lésions mélanocytaires les plus utilisés, on retrouve l'ABCD dermoscopique, l'analyse des patrons, la liste italienne en 7 points, la méthode de Menzies ou la liste en trois points. Ces différentes techniques permettent d'établir, à partir d'une liste de critères, la malignité ou bénignité de la lésion mélanocytaire (12,14,15).

L'algorithme ABCD de dermoscopie est l'un des premiers à avoir été développé, il est connu comme l'un des plus faciles à appliquer. Se fondant sur l'ABCDE décrit plus haut, à l'exception de la lettre E, il permet d'établir un score dermoscopique total (SDT) en associant à chaque critère ABCD un coefficient de multiplication. Il est considéré qu'un SDT inférieur ou égal à 4,75 désigne une lésion comme bénigne, un SDT supérieur à 5,45 comme maligne et celles se situant entre ces deux chiffres comme suspectes.(12)

			SCORE
Asymétrie	0,1 ou 2 axes : Couleur, contours, structure	0-2	Asymétrie x 1,3
Analyse des Bords	Arrêt abrupt de la pigmentation Lésion divisée en 8 segments	0-8	+ Analyse des Bords x 0,1
Couleur	De 1 à 6 couleurs	1-6	+ Couleur x 0,5
Structures Dermoscopiques	Réseau pigmentaire, aire sans structures, points, globules, stries	1-5	+ Structures dermoscopiques x 0,5
			= Score dermatoscopique total (SDT)

Tableau 1. Score dermatoscopique total calculé à partir des critères ABCD

La liste italienne en 7 points se compose de sept critères sémiologiques, dont trois majeurs et quatre mineurs. Elle permet une simplification de l'analyse des patrons et est considérée comme facilement utilisable. Un score est calculé à partir de ces 7 critères, toute lésion obtenant une note supérieure ou égale à 3 étant considérée comme un mélanome avec une sensibilité de 95% (12).

CRITÈRES DERMOSCOPIQUES	SCORE
Critères majeurs	
Réseau pigmenté atypique	2
Voile blanc-bleu	2
Architecture vasculaire atypique	2
Critères mineurs	
Stries irrégulières	1
Globules et points irréguliers	1
Tâches d'encre	1
Régression	1

Tableau 2. Critères dermoscopiques de la liste en 7 points

L'algorithme de Menzies s'adapte aux caractéristiques médicales et épidémiologiques de l'Australie, d'où provient son auteur, et est constitué de points positifs et négatifs. Pour qu'une lésion soit considérée comme un mélanome, il faut compter au moins un critère positif et aucun critère négatif. Cette méthode a pour but d'être très sensible et a été développée pour être utilisée par les médecins généralistes (12).

CARACTERISTIQUES POSITIVES	CARACTERISTIQUES NEGATIVES
<ul style="list-style-type: none"> • Voile blanc-bleu • Multiples points bruns agrégés • Pseudopodes • Stries rayonnantes • Zones de dépigmentation cicatricielle • Globules ou points noirs périphériques • Images en grain de poivre ou aspect poussiéreux • Polychromie marquée • Réseau pigmentaire élargi 	<ul style="list-style-type: none"> • Symétrie • Aspect monochrome

Tableau 3. Algorithme de Menzies

Si l'analyse des patrons architecturaux de dermoscopie est la stratégie la plus utilisée par les dermatologues, elle est cependant plus complexe à appliquer et a fait l'objet de nombreuses tentatives de simplification. Elle consiste dans un premier temps à identifier des patrons globaux puis, des patrons spécifiques (12,16).

Les patrons globaux sont : globulaire, pavimenteux, réticulaire, homogène, étoilé, parallèle, multicomposé et non spécifique (zones sans structures). Le patron multicomposé correspond à l'association de trois patrons dermoscopiques ou plus dans la même lésion, créant le plus souvent une forme asymétrique. Il correspond le plus souvent à celui du mélanome.

Les patrons spécifiques font référence à des structures particulières, de sémiologie élémentaire, à étudier sans prendre en compte les patrons globaux. Parmi ceux-ci, on retrouve : les stries, les pseudopodes, les

points noirs, le voile blanc-bleu, les zones blanches, les zones de régression, les tâches d'encre et le patron vasculaire (12,16).

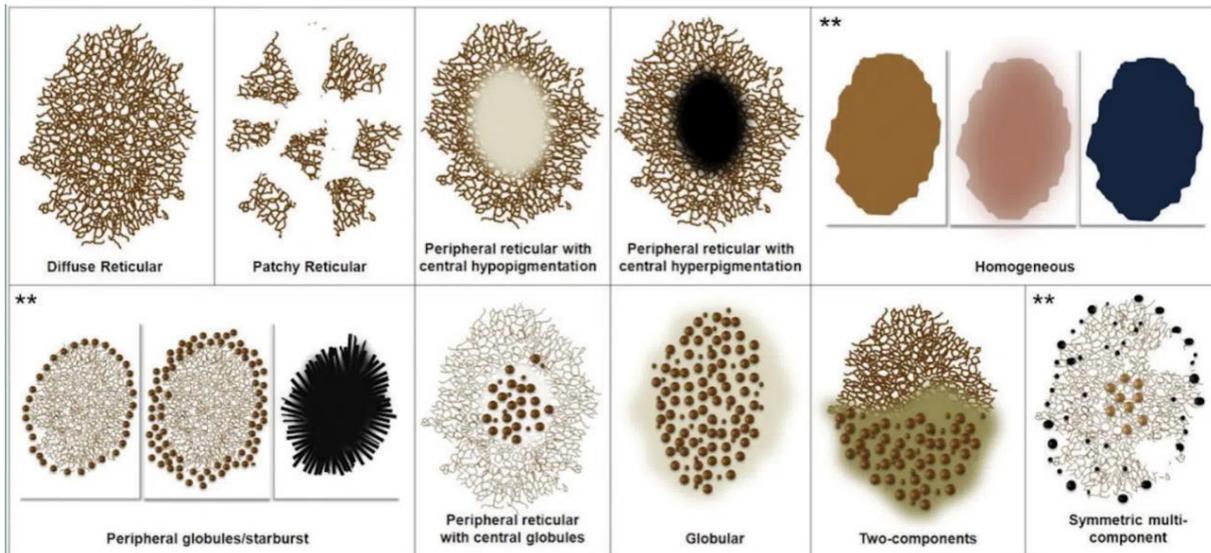


Figure 11 : Schématisation des différents patrons globaux selon l'analyse des patrons par Marghoob (16).

La règle des trois points est composée de trois critères : une asymétrie, un réseau atypique, présence de structures bleues-blanchâtres. La présence de deux de ces trois critères est évocatrice d'une lésion suspecte à risque d'être un mélanome. Cette méthode simple a été développée pour un usage par des médecins généralistes tout comme l'algorithme de Menzies (12).

La dermoscopie opère ainsi par l'application d'algorithmes et se base sur la reconnaissance de patrons visuels, permettant aux dermatologues entraînés de classer les lésions cutanées. L'analyse d'images par le biais d'algorithmes est un processus particulièrement courant dans l'utilisation de l'intelligence artificielle, qui a connu, ces dernières années, une expansion dans des domaines très variés (10). Son application en dermoscopie est un sujet de recherche de grande prévalence, ce qui nous a motivées à nous pencher sur l'état actuel de la science à ce sujet.

I-3 PRINCIPE DE L'INTELLIGENCE ARTIFICIELLE DANS L'ANALYSE DES LÉSIONS CUTANÉES PIGMENTÉES :

Dans une multitude de domaines et de disciplines, l'intelligence artificielle (IA) et l'apprentissage automatique (AA ou Machine Learning) peuvent aider les cliniciens à prendre de meilleures décisions cliniques, voire remplacer, dans certaines situations, le diagnostic du praticien. Les algorithmes basés sur l'apprentissage profond (AP ou Deep Learning) et sur les réseaux de neurones convolutifs (ConvNets ou CNN pour Convolutional Neural Network) (définitions ci-dessous) peuvent considérablement améliorer l'efficacité de la classification d'images, de la détection d'objets, de la segmentation etc...(17). La capacité d'apprentissage exponentielle de l'IA lui permet de percevoir des différences subtiles dans les caractéristiques d'une image telles que la taille, la texture, les différentes teintes/couleurs...et peut surpasser les capacités visuelles et analytiques de l'humain (18).

Depuis de nombreuses années, dans le domaine de la dermatologie, la tendance à la numérisation des images en dermatoscopie et à l'utilisation de ces nouvelles technologies dans l'aide au diagnostic est en pleine expansion (19,20). Pour ce travail de thèse, nous souhaitons examiner, sur les deux dernières années et de façon systématique, les études récentes portant sur les algorithmes d'IA et d'AA, visant à faciliter le diagnostic précoce des cancers de la peau, qui mettent l'accent sur leur utilisation possible en soins primaires. Pour ces raisons, il nous semblait nécessaire de définir l'IA et les principes fondamentaux de ces technologies avant de pouvoir discuter de leur efficacité, précision et sûreté dans l'analyse de lésions cutanées.

L'IA fait référence à l'intelligence exprimée par des machines (elles-mêmes créées par l'Homme) qui simulent les processus cognitifs humains tels que l'apprentissage et la résolution de problèmes (21). Les « agents intelligents » sont des systèmes capables d'observer et d'apprendre de leur environnement, mais aussi d'agir et d'utiliser des connaissances pour atteindre un objectif, réaliser des tâches, tout en étant capable de s'adapter (21).

En 1943, le premier réseau neuronal artificiel (ANN pour Artificial Neural Network) composé de circuits électriques a été modélisé pour simuler les interactions interneuronales du cerveau humain (22). Ces réseaux de neurones sont une composante fondamentale de l'apprentissage automatique (AA) et se trouvent au cœur des algorithmes de l'apprentissage en profondeur (Deep Learning). Leur appellation et leur architecture sont inspirées du fonctionnement du cerveau humain, mimant le mode de communication biologique entre les neurones (notamment les synapses).

Un ANN est composé de plusieurs couches nodales, comprenant une couche d'entrée pour recevoir les données, une ou plusieurs couches cachées où se déroule leur traitement et une couche de sortie qui donne le résultat (23). Chaque couche est composée de neurones artificiels, également appelés nœuds ou unités. Chaque neurone reçoit des entrées pondérées depuis les neurones de la couche précédente, calcule une somme pondérée de ces entrées, puis applique une fonction d'activation pour produire une sortie. Chaque nœud, ou neurone artificiel, est connecté à d'autres et possède un poids et un seuil associé. Ces poids déterminent l'importance relative des différentes entrées pour chaque neurone. De plus, chaque neurone possède un seuil qui influence sa sensibilité à l'activation. Si la sortie d'un nœud dépasse le seuil spécifié, celui-ci s'active et transmet des données à la couche suivante du réseau. Dans le cas contraire, aucune donnée n'est transférée vers la couche suivante. Ainsi, schématiquement, le neurone artificiel est modélisé à partir du neurone biologique. Il possède un comportement binaire qui fonctionne comme le flux neuronal (pour simplifier : les dendrites reçoivent les signaux, le corps cellulaire les intègre, l'axone les transmet à d'autres neurones et les synapses permettent la communication entre les neurones en libérant des neurotransmetteurs) :

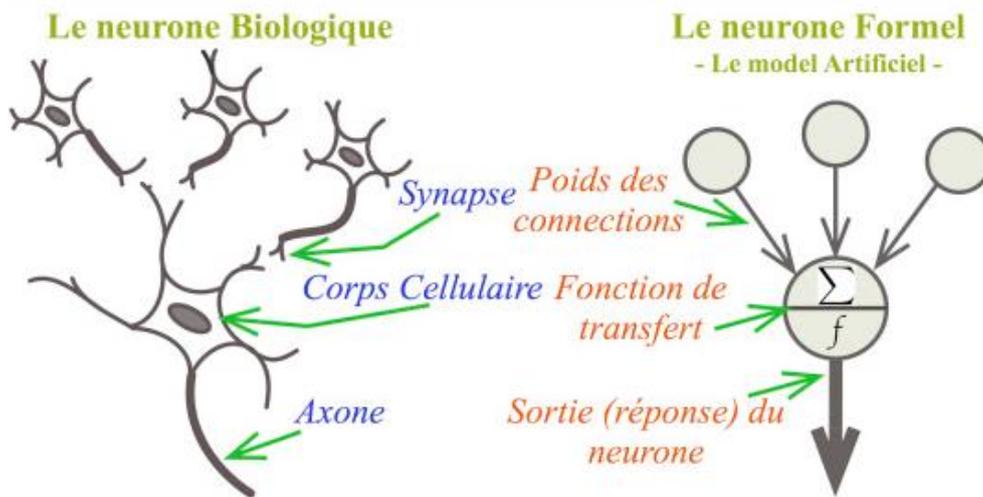


Figure 12 : Neurone artificiel simulant le comportement d'un neurone biologique d'après Zemouri *et al.* (24).

Après avoir produit une sortie, le réseau la compare à celle désirée (si disponible). En utilisant un algorithme d'apprentissage tel que la rétropropagation du gradient, le réseau ajuste ses poids et seuils afin de minimiser l'écart entre les sorties réelles et les sorties désirées (23). Ce processus itératif permet au réseau de s'entraîner sur des données d'apprentissage et d'améliorer ses performances au fil du temps. Les performances d'un ANN dépendent du nombre et de la structure de ses couches neuronales et de l'ensemble des données d'entraînement qui ont été entrées (25).

Le concept d'IA est né quelques années après le premier ANN, en 1956, au Dartmouth College dans le New Hampshire (26). Trois ans plus tard, la première recherche informatique utilisant un réseau de neurones artificiels est réalisée (27). En 1963, l'aide au diagnostic par ordinateur a été appliquée pour la première fois dans l'analyse de nodules pulmonaires à partir de radiographies thoraciques (28).

L'apprentissage automatique (Machine Learning) et l'apprentissage profond (Deep Learning) sont deux sous-domaines de l'IA qui diffèrent par leurs approches, leurs architectures et leurs capacités. L'AA est une méthode d'enseignement aux ordinateurs qui leur permet de réaliser des tâches en se basant sur des données et des algorithmes (régression linéaire, arbres de décisions, méthodes de clustering...) sans être explicitement programmées pour chaque étape (29). En général, les algorithmes d'apprentissage automatique sont utilisés pour faire une prédiction ou une classification et ils prennent des décisions sans programmation préalable (23). Sur la base de certaines données d'entrées, qui peuvent être étiquetées (données pour lesquelles la réponse cible est connue) ou non, l'algorithme produira une estimation d'un modèle. Puis, une fonction d'erreur évalue la prédiction de ce modèle. Celle-ci, nommée également fonction de perte, ou encore fonction objective, est une mesure utilisée pour évaluer à quel point les prédictions d'un modèle correspondent aux valeurs réelles des données d'entraînement. En effet, s'il existe des exemples connus, une fonction d'erreur peut effectuer une comparaison pour en évaluer la précision. Enfin, grâce à un

processus d'optimisation du modèle, les poids sont ajustés pour réduire l'écart entre l'exemple connu et l'estimation générée. L'algorithme répétera ce processus itératif "d'évaluation et d'optimisation" en mettant à jour les poids de manière autonome jusqu'à ce qu'un seuil de précision soit atteint. L'apprentissage automatique repose sur des méthodes traditionnelles telles que les machines à vecteurs de support (SVM), les arbres de décision, les réseaux bayésiens, etc.

Le principe des machines à vecteurs de support est basé sur la recherche d'un hyperplan optimal (concept géométrique analytique décrivant un sous-espace de dimension) qui sépare les données en classes différentes tout en maximisant la marge entre elles, de manière à pouvoir placer les nouveaux points de données dans la bonne classe à l'avenir (figure 13 d'après (30)).

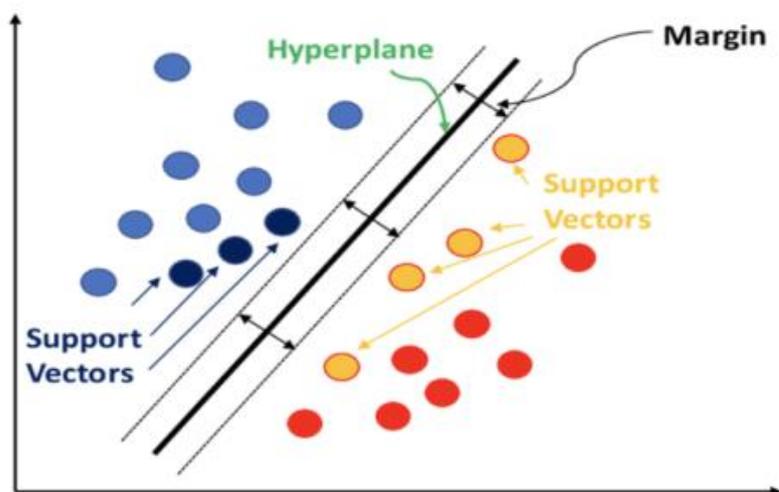


Figure 13 : Principe d'une machine à vecteur de support.

- **Marge** : l'écart entre l'hyperplan et les vecteurs de support.
- **Hyperplan** : frontières de décision qui aident à classer les points de données.
- **Vecteurs de support** : points de données qui se trouvent sur ou à proximité de l'hyperplan et influencent la position de l'hyperplan.
- **Fonction noyau** : fonction utilisée pour déterminer la forme de l'hyperplan et de la frontière de décision.

Dans l'étude de Melbin *et al* (31), les caractéristiques «ABCD», pour « Asymétrie Bordure Couleur Diamètre », d'une lésion cutanée sont utilisées pour extraire des données telles que la forme, la couleur et la taille de chaque image clinique fournie. Après avoir sélectionné ces caractéristiques, la lésion cutanée est classée à l'aide de SVM en mélanome, kératose séborrhéique et lupus érythémateux (figure 14). Cette méthode d'utilisation de l'ABCD dans des SVM donne d'excellents résultats tout en fournissant des informations significatives (31).

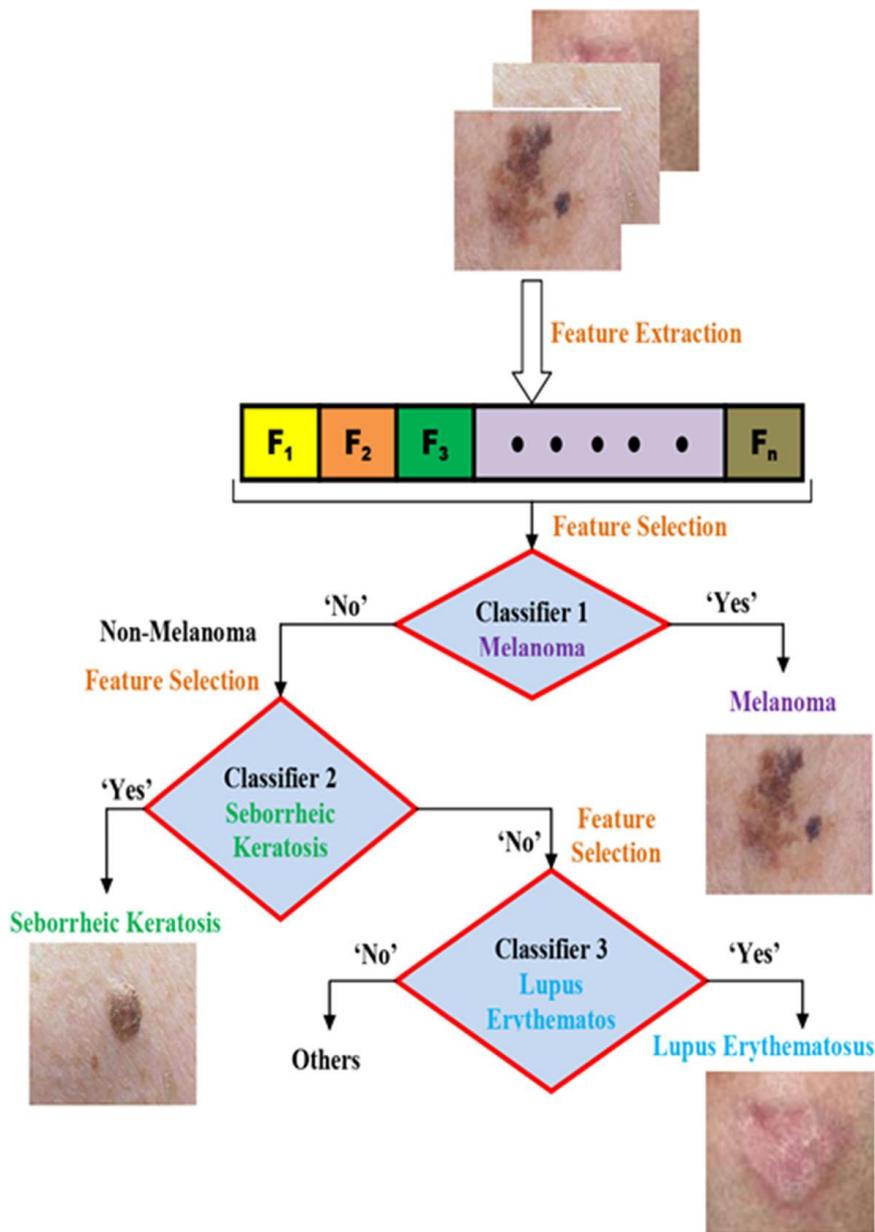


Figure 14 : Organigramme de la classification de trois maladies de la peau par SVM d'après Melbin et al.

L'image de la lésion cutanée est traitée, et des caractéristiques telles que la texture, la couleur et la forme sont extraites. Ces caractéristiques sont représentées par F_1, F_2, \dots, F_n . L'objectif principal de la classification par SVM est de trouver l'hyperplan qui maximise la marge entre deux classes.

Les SVM sont actuellement utilisés en tant qu'aide au diagnostic assisté par ordinateur pour classer les lésions cutanées en mélanomes ou non mélanomes (figure 14) (32).

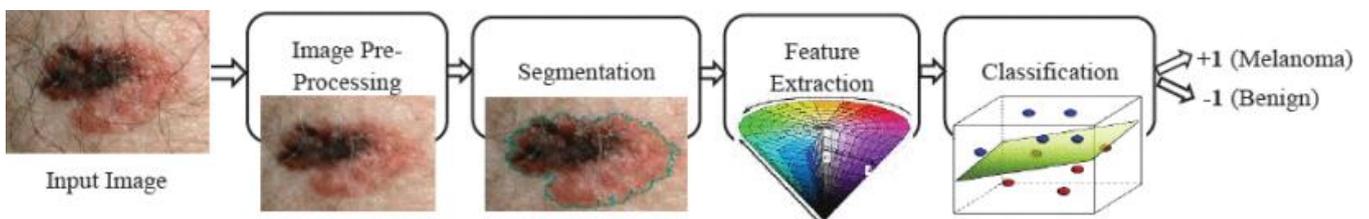


Figure 15 : La structure du système CAD (Computer-Aided Design ou Conception Assistée par Ordinateur) de l'étude d' Afifi et al. (33). Le classificateur SVM a été testé pour la détection du mélanome et les résultats expérimentaux ont confirmé une grande précision dans la classification.

Le principe de l'arbre de décision est une technique d'apprentissage automatique utilisée pour résoudre des problèmes de classification et de régression. L'idée fondamentale consiste à diviser récursivement l'ensemble des données en sous-ensembles plus petits en fonction des caractéristiques des données, jusqu'à ce qu'ils soient suffisamment purs en termes de classe ou qu'une condition d'arrêt soit atteinte (34). A titre d'exemple,

en dermatologie, les arbres de décision ont été utilisés pour créer un algorithme intuitif aidant à quantifier le risque à long terme de cancer de la peau (hors mélanome) après une greffe hépatique (35).

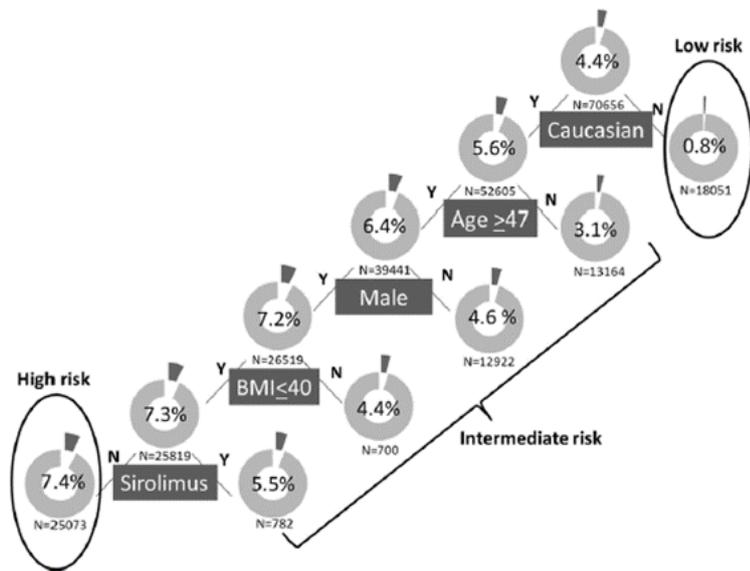


Figure 16 : Modèle d'arbre de décision développé dans Tanaka et al. (35). Les rectangles indiquent les facteurs utilisés et les valeurs seuils pour différencier les patients. Les diagrammes circulaires indiquent le taux d'incidence du cancer de la peau non-mélanome (NMSC) pour chaque groupe. Ces derniers ont été classés en trois sous-groupes (risques élevé, intermédiaire et bas) en fonction du taux d'incidence.

Les réseaux bayésiens sont fondés sur le théorème de Bayes (calcul de la probabilité postérieure d'un événement en se basant sur d'autres événements vérifiés ou mise à jour de nos croyances initiales sur un événement lorsque l'on obtient de nouvelles informations)(36). Ce principe est déjà utilisé pour détecter les maladies de peau, pour la segmentation (partitionnement d'une image en plusieurs régions ou segments distincts) et classification des images cliniques en dermatologie avec une précision satisfaisante (37,38). Ce processus est réalisé de manière itérative, ce qui, comparativement à d'autres méthodes, demande moins de ressources informatiques et de sessions d'apprentissage.

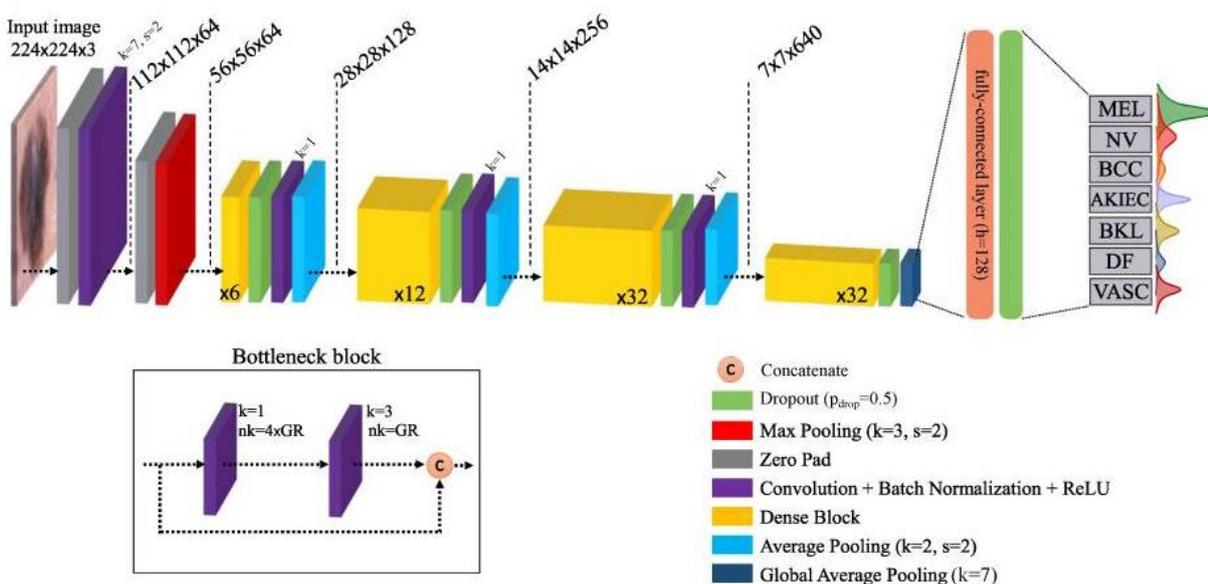


Figure 17 : Schéma de l'architecture du modèle Bayésien DenseNet-169 d'après Mobiny et al (39).

L'ensemble des différentes étapes séquentielles du système figure sur ce diagramme. Sans donner de réponse unique « c'est un mélanome ou c'est un naevus », le réseau bayésien produit une distribution prédictive par classe (MEL mélanome, NV naevus, BCC carcinome baso-cellulaire, AKIEC Kératoses actiniques et carcinome intra-épithélial, BKL Kératose bénigne, DF dermatofibrome et VASC lésions vasculaire) dont la moyenne et la dispersion représentent respectivement la prédiction et l'incertitude du réseau.

Les notions d'AA et d'apprentissage profond ont tendance à être utilisées de manière interchangeable mais il convient de noter des nuances entre les deux. L'AP est en fait une sous-catégorie de l'AA, fondée uniquement sur les ANN composés de nombreuses couches (d'où le terme "profond") et sur l'apprentissage par représentation. La différence entre l'apprentissage profond et l'apprentissage automatique réside dans la manière dont chaque algorithme apprend. Tout comme l'AA, l'apprentissage automatique "profond" peut utiliser des ensembles de données étiquetées, également connus sous le nom d'apprentissage supervisé, pour informer son algorithme, mais il n'en a pas nécessairement besoin. Le processus d'apprentissage profond peut ingérer des données non structurées sous leur forme brute (textes, images...) et déterminer automatiquement l'ensemble des caractéristiques qui distinguent ces différentes catégories de données. Cela limite l'intervention humaine et permet d'utiliser de très grandes quantités de données (23).

Les principaux réseaux d'apprentissage profond existants sont :

- les réseaux neuronaux récurrents (RvNN, Recursive Neural Networks) ;
- les réseaux neuronaux récurrents (RNN, Recurrent Neural Networks) ;
- les réseaux neuronaux auto-organiseurs de Kohonen (KNN, k-nearest neighbors algorithm) ;
- les réseaux neuronaux adversaires génératifs (GAN, Generative Adversarial Neural) ;
- les réseaux neuronaux convolutifs (CNN).

Les RvNN sont des architectures de neurones qui s'appliquent à des structures de données hiérarchiques et récursives (capacité d'un algorithme à s'appeler lui-même de manière répétée pour résoudre un problème), telles que des arbres syntaxiques dans le traitement du langage naturel ou des structures d'analyse syntaxique (40). Contrairement aux RNN, les réseaux neuronaux récurrents sont conçus pour traiter des données structurées de manière arborescente plutôt que séquentielle. Ils en exploitent la récursivité pour en extraire des informations hiérarchiques aux différentes profondeurs de l'arborescence.

Les RNN sont un type spécifique de réseau de neurones dans lesquels les connexions inter-neuronales forment une boucle (connexions récurrentes) qui leur permettent de traiter des données séquentielles en prenant en compte les états précédents. Ils sont couramment utilisés pour les séquences de données telles que du texte, des séries temporelles ou des signaux audios. Ils ont déjà été utilisés pour la détection de lésion de mélanome (41). Dans cette étude, la phase de classification du modèle proposé emploie des techniques d'apprentissage profond en combinant la notion d'optimisation dans un RNN (41).

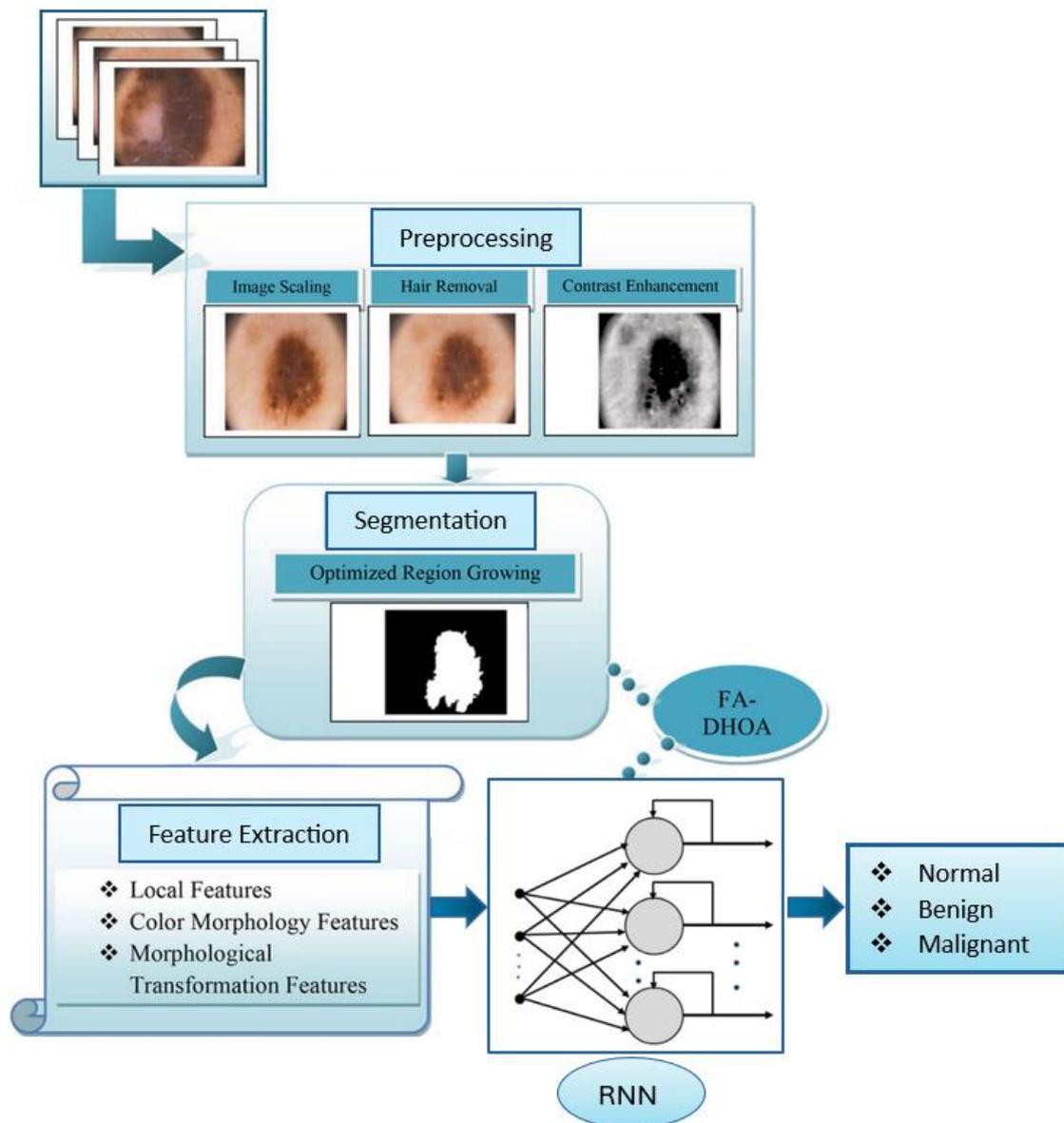


Figure 18 : Représentation architecturale globale du modèle de détection de mélanome proposé dans Divya *et al.* et architecture du RNN d'après Abu Zitar *et al.*

Le modèle proposé pour la détection du cancer de la peau comprend quatre étapes principales : prétraitement, segmentation, extraction des caractéristiques et classification. Les images dermoscopiques sont d'abord soumises à une étape de prétraitement qui inclut l'amélioration de l'image et la suppression des poils. Puis la segmentation de la lésion est effectuée par un algorithme optimisé de croissance régionale. Lors de la phase d'extraction des caractéristiques locales, de morphologie et de couleur, sont extraites. La phase de classification utilise un algorithme d'apprentissage profond modifié en intégrant un concept d'optimisation dans un réseau de neurones récurrents. Les performances de la croissance régionale et du RNN sont améliorées par l'algorithme d'optimisation de la chasse au cerf modifié (Deer Hunting Optimization Algorithm DHOA) appelé Algorithme DHOA Adaptatif de Fitness (FA-DHOA).

Les réseaux de Kohonen sont utilisés pour réaliser un type de clustering ou de cartographie de données. L'objectif principal d'un KNN est de projeter les données d'entrée dans un espace de dimension réduite tout en préservant les relations topologiques entre elles. Les KNN apprennent de manière non supervisée en ajustant les poids des neurones de sortie afin que les neurones similaires (correspondant à des données similaires) se trouvent à proximité les uns des autres sur la carte. Ils ont de nombreuses applications et peuvent être utilisés pour l'analyse de données géospatiales telles que les cartes climatiques, les cartes de

population, les images satellites ou encore être appliqués à la segmentation d'images pour identifier des régions d'intérêt ou des structures. Les réseaux de Kohonen ont été appliqués pour détecter la croissance anormale de lésions cutanées (42) ou encore pour classer et distinguer des mélanomes de lésions cutanées bénignes (43).

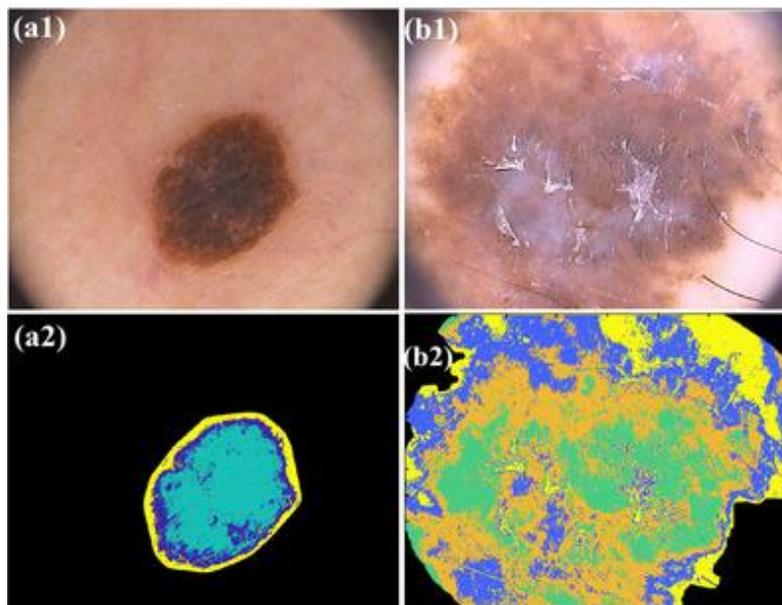


Figure 19 : Segmentation de lésions grâce au KNN. Une couleur donnée n'est considérée comme présente que si elle occupe une surface supérieure à 5 % de la surface de la lésion. Le KNN a segmenté la lésion **(a)** en 3 classes (ou couleurs) et la lésion **(b)** en 4 classes.

Les GAN sont utilisés pour générer de nouvelles données réalistes à partir d'un ensemble de données d'apprentissage, sans avoir besoin d'en connaître la distribution de probabilité sous-jacente (44). Ils se composent de deux réseaux neuronaux profonds : un générateur et un discriminateur en compétition lors de l'entraînement. Le générateur apprend à produire de nouvelles données à partir de données aléatoires et tente de créer des données qui ne se distinguent pas de celles qui sont réelles. Le discriminateur apprend à les distinguer des données réelles provenant de l'ensemble de données d'apprentissage. Au fur et à mesure de l'entraînement, le générateur et le discriminateur s'améliorent, le premier en essayant perpétuellement de tromper le discriminateur, le second en progressant dans sa capacité à distinguer les données générées des données réelles. Ce processus est également largement utilisé en dermatologie ; le véritable potentiel de l'AP n'est pas exploité à 100 % en raison de la taille limitée des données d'entraînement. Les GAN permettent de générer de nouvelles données d'apprentissage afin d'augmenter les ensembles d'entraînements existants, à savoir le nombre d'images dans les bases de données, et d'accroître les performances de réseaux neuronaux tels que les CNN (45,46).

Les réseaux de neurones convolutifs, également appelés réseaux de neurones convolutionnels, sont le sous-type d'ANN le plus souvent utilisé pour le traitement et la détection d'images en médecine et plus particulièrement en dermatologie, en anatomopathologie et en radiologie (47). En effet, le CNN est un type

de réseau de neurones artificiels spécialement bien adapté à la tâche de reconnaissance de motifs dans des données structurées en « grilles » telles que les images, les séquences temporelles (collection de points de données organisés de manière séquentielle et indexés par des instants temporels spécifiques) et les données volumétriques (données tridimensionnelles).

Ci-dessous, les principaux composants et le fonctionnement d'un réseau de neurones convolutionnels (schématisés en figure 20) (48):

- Couche de convolution : c'est la pierre angulaire d'un CNN. Elle est composée de filtres (ou noyaux) qui sont appliqués à des régions locales de l'entrée (en l'occurrence ici l'image à traiter) en effectuant une opération de convolution. Celle-ci consiste à multiplier les valeurs de l'entrée par les poids du filtre, puis à totaliser les résultats pour produire une carte d'activation. La couche de convolution sert à extraire les caractéristiques importantes des données d'entrée (contours, textures ou motifs...).
- Couche de pooling : après la couche de convolution, cette couche est généralement ajoutée pour réduire la dimension spatiale de la représentation et le nombre de paramètres du modèle. Elle réduit la taille, en largeur et hauteur, des cartes d'activations générées par les couches de convolution en agrégeant les informations à partir de régions voisines tout en préservant les données essentielles.
- Activation non linéaire : entre chaque couche de convolution ou de pooling, cette fonction est appliquée aux cartes d'activation pour introduire de la non-linéarité dans le modèle (permettant au réseau de neurones d'apprendre des fonctions de décisions complexes). La plus couramment utilisée est la fonction ReLU (Rectified Linear Unit), mais d'autres, telles que la sigmoïde ou la tangente hyperbolique peuvent également être appliquées.
- Réseaux entièrement connectés : après plusieurs couches de convolution et de pooling, les informations sont aplaties (flattening) et introduites dans des couches entièrement connectées, composées de neurones traditionnels. Ces couches finales sont utilisées pour effectuer la classification ou la régression en produisant des prédictions sur les classes ou valeurs cibles.
- Entraînement : les CNN sont généralement entraînés à l'aide d'un algorithme d'optimisation tel que la descente de gradient stochastique ou l'une de ses variantes (contrairement à la descente de gradient classique, le gradient stochastique se calcule sur un seul exemple d'entraînement ou un petit sous-ensemble aléatoire d'exemples. L'algorithme devient ainsi plus rapide et plus évolutif car il n'a pas besoin de traiter l'intégralité des données à chaque étape). L'entraînement s'effectue en ajustant les poids des filtres du réseau afin de minimiser une fonction de perte, souvent associée à la tâche spécifique telle que la classification ou la régression.

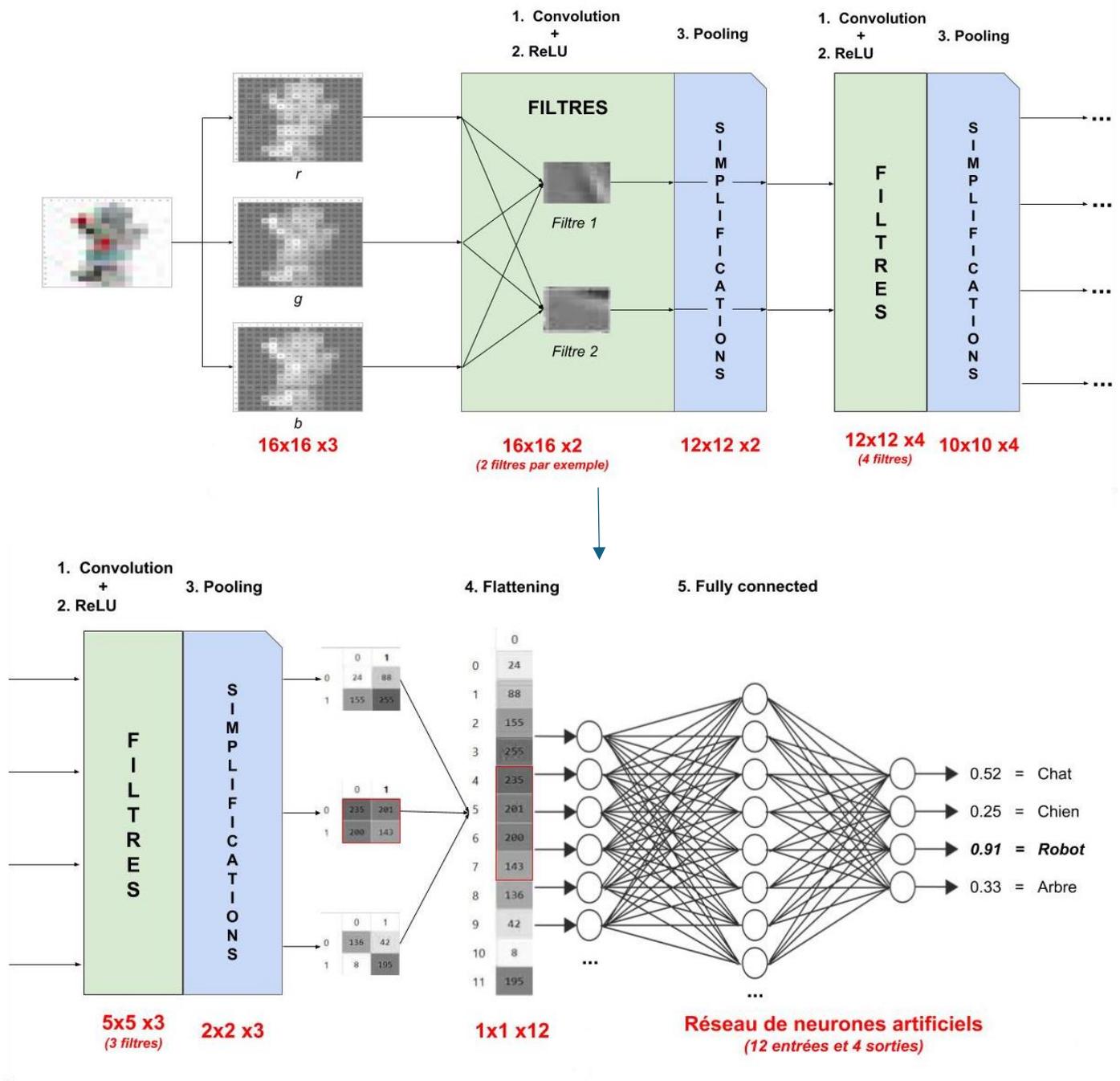


Figure 20 : Extraction des informations de l'image grâce à un enchaînement de filtres (et de simplifications) puis prédiction de la classe de l'image. D'après Lambert R. (48), le réseau de neurones convolutionnels va recevoir cette image (vue comme une matrice de nombres organisés) et appliquer un certain nombre d'opérations divisées en deux catégories :

- les filtres, pour mettre en évidence, par exemple, les arêtes verticales, les variations de coloration, etc... ; plus ils tardent à intervenir dans l'enchaînement, plus efficaces ils sont dans la détection des formes complexes ;
- les simplifications pour alléger les calculs et en faire ressortir les informations principales.

Ensuite l'ANN proposera comme hypothèse que l'image représente un robot, un chien, un humain...

Actuellement, les architectures CNN les plus utilisées dans le domaine de la dermatologie sont *GoogLeNet*, *Inception-V3, V4*, *ResNet*, *Inception-ResNet V2* et *Dense Net* (49). Pour qu'elles aient une bonne précision de diagnostic et un algorithme de sensibilité et de spécificité élevé, elles ont besoin d'un grand nombre d'image

de haute qualité. Les ensembles d'images les plus couramment utilisés aujourd'hui pour entraîner les systèmes d'IA en dermatologie sont les archives ISIC (2016-2021), HAM10000, BCN20000 et l'ensemble d'images PH².

La base de données ISIC (International Skin Imaging Collaboration) est une collection de milliers d'images dermatoscopiques représentant des images de diverses lésions cutanées, allant de lésions bénignes aux mélanomes, annotées par des dermatologues experts. Le principal objectif consiste à fournir une ressource standardisée pour entraîner et valider des modèles d'IA dans le diagnostic automatisé des maladies de la peau. Utilisée dans des compétitions comme le ISIC Challenge, elle contribue à améliorer la performance des algorithmes d'IA dans la détection de cancers cutanés. La diversité des patients et des types de peau constituant cette base de données permet de tester l'efficacité des modèles dans différents contextes cliniques (50).

La base de données HAM10000 (Human Against Machine with 10000 training images) est un ensemble de données dermatoscopiques contenant 10 015 images de lésions cutanées provenant de plusieurs sources et couvre 7 types de lésions communes : kératoses actiniques (AKIEC), carcinomes basocellulaires (BCC), lésions kératosiques bénignes (BKL), dermatofibromes (DF), mélanomes (MEL), naevus mélanocytaires (NV) (grains de beauté) et lésions vasculaires (VASC). Elles ont été annotées par des dermatologues confirmés, ce qui en fait une ressource précieuse pour l'entraînement et l'évaluation de modèles d'IA (51).

La base de données BCN20000 est une collection de 18 946 images dermatoscopiques collectées entre 2010 et 2016 à l'Hôpital Clínic de Barcelone (BCN) qui couvre 5 583 lésions cutanées. Elle comporte des images couvrant 8 catégories diagnostiques principales, incluant les naevus, mélanomes, carcinomes et kératoses, ainsi qu'une catégorie supplémentaire pour les lésions non classifiables (OOD pour out-of-distribution). Elle inclut des cas cliniquement complexes, comme des lésions situées dans des zones difficiles (ongles, muqueuses) ou des lésions hypopigmentées et de grande taille. Les images sont accompagnées de métadonnées détaillées (localisation, âge, sexe, date de capture) et les diagnostics de mélanomes sont validés par histopathologie (52).

La base de données PH² contient 200 images dermatoscopiques haute résolution, chacune représentant une lésion cutanée, telles que des naevus bénins, des naevus dysplasiques et des mélanomes. Des dermatologues experts ont annoté chacune des images et indiqué les contours manuels des lésions ainsi que des caractéristiques dermatoscopiques spécifiques comme les structures pigmentaires, les régions homogènes et les régions de couleur irrégulière (53).

Dans la revue systématique de la littérature d'O.T. Jones *et al.*, les auteurs se sont intéressés à 272 articles publiés du 1^{er} janvier 2000 au 9 août 2021, afin d'évaluer la précision diagnostique des algorithmes d'IA/AA dans la détection des cancers de la peau en soins primaires. Ils ont relevé les points suivants :

- la précision diagnostique moyenne était plutôt raisonnable (89,5 % [IC 59,7-100 %] pour le mélanome, 85,3 % [71-97,8 %] pour le carcinome épidermoïde et 87,6 % [70-99,7 %] pour le carcinome basocellulaire) ;
- les ensembles de données utilisés pour l'AA provenaient, cependant, en grande partie de patients recrutés dans des établissements cliniques spécialisés (et non pas de populations à faible prévalence de cancer cutané comme retrouvées en soins primaires) ;
- il existait, en outre, une grande disparité dans l'origine et le type de données utilisées pour former et tester les algorithmes d'IA/AA (l'origine de l'image, les critères d'inclusion et les données démographiques du patient (y compris la couleur et le type de peau) ne sont pas rapportés de manière exhaustive).

Les algorithmes d'IA/AA étaient alors développés et testés sur des données qui n'étaient pas représentatives de l'ensemble de la population pour laquelle ils étaient destinés à être utilisés et cela entraînait forcément un biais inhérent.

L'objectif principal de notre travail était d'évaluer si, à ce jour, l'IA pouvait être utilisée afin d'aider les médecins généralistes à établir un diagnostic et à prendre en charge précocement des lésions malignes dermatologiques. Notre objectif secondaire était de recenser les différentes méthodes d'IA existantes, en déterminer le coût, l'efficacité et l'accessibilité par rapport à l'intervention de spécialistes.

II - MATERIEL ET MÉTHODE :

II-1 TYPE D'ÉTUDE :

Afin d'atteindre notre objectif, nous avons effectué une revue systématique de la littérature en suivant les critères définis par les recommandations internationales « Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) ». Cette recherche, menée par deux chercheurs, a été réalisée à l'aide de la méthode décrite dans l'article de Dupouy *et al.* publié en 2013 (54).

II-2 SOURCES D'INFORMATIONS ET STRATÉGIE DE RECHERCHE :

Préalablement au travail de recherche, nous avons parcouru la base de données SUDOC (Système Universitaire de Documentation) afin d'identifier les travaux de thèse existant, en lien avec le sujet, vérifiant ainsi l'existence éventuelle d'un travail de recherche trop comparable. Nous avons également exploré la littérature grise en consultant les sites d'agences de santé nationales et internationales (la Haute Autorité de Santé (HAS), l'Institut National du Cancer (INCa), l'Organisation Mondiale de la Santé (OMS), etc...).

Pour effectuer nos recherches, nous avons interrogé trois bases de données bibliographiques : MEDLINE, Embase et Web of Science en nous appuyant sur les équations de recherches utilisées dans la revue systématique de la littérature d'OT Jones et al. (9) du journal the Lancet. Celle-ci utilisait des mots clés relatifs à l'IA/AA, au cancer de la peau, à la détection précoce et aux soins primaires (après les avoir traduits en anglais et utilisés le thésaurus MeSH pour l'indexation et la normalisation des termes recherchés). Ces équations de recherches sont répertoriées dans l'annexe 1.

II-3 CRITÈRES DE SÉLECTION ET D'EXCLUSION :

Critères de sélection des articles :

- articles de recherche publiés dans des revues à comité de lecture ;
- méthodologie des articles : études observationnelles, essais cliniques randomisés et non randomisés ;
- langue : anglais ou français ;
- date de publication : du 1^{er} août 2021 au 31 août 2023 ;
- études utilisant des données provenant de population de soins primaires et secondaires (patients ayant déjà été évalués et orientés) ;
- études fournissant des preuves de la précision, de l'utilité et de l'acceptabilité de l'application d'algorithmes d'IA/AA au triage ou au diagnostic du cancer de la peau.

Critères d'exclusion des articles :

- articles non publiés en anglais ou en français ;
- articles publiés hors de la période définie ;
- revues systématiques de la littérature et Méta-Revues ;

- études décrivant uniquement le développement d'un algorithme d'IA/AA, sans aucun test ou évaluation effectués sur une population donnée ou un ensemble d'image donné ;
- études n'incorporant pas d'élément d'apprentissage automatique (ne contenant ni étapes d'entraînement, ni test, ni validation) ;
- études utilisant l'IA/AA dans la découverte de biomarqueurs, dans l'analyse d'image d'histopathologie et dans l'identification de protéines cancéreuses par spectrométrie de masse ;
- études dont la taille des échantillons était inférieure à 50 patients (photos) atteints d'un cancer de la peau et à 50 patients dans le groupe contrôle ;
- études dont le facteur d'impact était inconnu ou inférieur à 1.

II-4 SÉLECTION DES ÉTUDES :

La sélection des articles s'est effectuée en trois étapes principales :

- une première analyse par lecture simple des titres en excluant les revues systématiques de la littérature ou méta revue, les articles rédigés dans une langue autre qu'anglais ou français, les titres hors sujet ;
- une seconde analyse portant sur les résumés des articles avec intégration des plus pertinents pour répondre au sujet de recherche et éviction des publications répondant à un ou plusieurs critères d'exclusions ;
- une troisième analyse par lecture intégrale de l'ensemble des articles restants et sélection de ceux remplissant les critères d'inclusions cités ci-dessus.

Au niveau des deux premières étapes, chacun des deux auteurs a évalué la moitié des textes sélectionnés. Puis, les articles dont la classification semblait ambiguë ont fait l'objet d'une concertation. La troisième étape a consisté, pour les deux chercheurs, à lire l'intégralité des articles répertoriés puis à mettre en commun les résultats de leurs analyses. Les articles dont l'inclusion ou l'exclusion était discordante ont fait l'objet d'une discussion commune et, si une incertitude subsistait, l'arbitrage final revenait au directeur de thèse.

II-5 ÉVALUATION DES ARTICLES SÉLECTIONNÉS :

Comme préconisé dans l'article du Pr. Dupouy (54), chaque étude scientifique sélectionnée a été évaluée à l'aide d'échelles et de grilles d'évaluation standardisées : afin de juger de la qualité et de la transparence du rapport en tant que tel d'une étude, nous avons utilisé la grille CONSORT (Consolidated Standards of Reporting Trials) (55) pour les essais cliniques randomisés, TREND (Transparent Reporting of Evaluations with Nonrandomized Designs) (56) pour les non randomisés et la grille STROBE (Strengthening The Reporting of Observational Studies in Epidemiology) (57) pour les études observationnelles (l'échelle STROBE aborde les trois principaux types d'études observationnelles : les études de cohorte, les études cas-témoins et les études transversales).

La qualité méthodologique de ces études a été évaluée à l'aide du score de Jadad (58) pour les essais cliniques randomisés et de l'échelle MINORS (Methodological index for non-randomized studies) (59) pour les non randomisés. Dans notre analyse de la littérature, les études observationnelles sélectionnées portaient sur l'évaluation de méthodes diagnostiques. Étant donné que l'échelle Newcastle-Ottawa (60,61) est principalement conçue pour comparer des cohortes exposées et non exposées, l'échelle QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies) (59,60) s'est imposée comme le choix le plus pertinent puisqu'elle vise à évaluer la qualité méthodologique des études de précision diagnostique. Cette dernière évalue quatre domaines spécifiques : la sélection des patients, le test étudié, le test de référence ainsi que le déroulement et la temporalité de l'étude. Pour chaque item, une réponse est donnée sous forme de « oui », « non », ou « incertain ». Afin de faciliter la classification et l'évaluation des études sélectionnées, nous avons attribué un point (1) à chaque item lorsqu'il est présent et zéro (0) lorsqu'il est absent (voir annexe 4).

Pour guider l'évaluation des études identifiées dans cette revue (éligibilité, phase de développement), nous nous sommes également appuyées sur la méthodologie de la revue systématique de la littérature O T Jones *et al.* parue dans le Lancet en juin 2022 (9). Il s'agissait d'utiliser le « CanTest Framework » (62), cadre élaboré en 2019, pour l'évaluation des tests et des approches diagnostiques.

Au final, coter notre propre revue avec la grille PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (63) nous a permis d'en estimer la rigueur de la rédaction et l'échelle AMSTAR (A Measurement Tool to Assess Systematic Reviews) (64) la qualité méthodologique.

III - RÉSULTATS :

III - 1 SÉLECTION DES ARTICLES :

Au total, 2 339 articles ont été identifiés à partir de trois bases de données (les équations de recherche se trouvent en annexe 1) : Web of Science (801 articles), Embase (570 articles) et PubMed (968 articles). Après avoir limité les publications aux années 2021 à 2023, nous avons obtenu 741 articles sur Web of Science, 499 sur Embase et 806 sur PubMed, soit un total de 2 046 articles. Après suppression des doublons (617 articles), il en restait 1 429.

La lecture des titres a permis d'en exclure 817 (hors sujets, revues de la littérature, méta-analyses, articles non rédigés en anglais ou en français ou articles hors-période/non disponibles), réduisant ainsi leur nombre à 612. Après l'examen des résumés, 53 articles ont été retenus, éliminant ceux qui ne sont pas axés sur les soins primaires ou les données cliniques. 6 d'entre eux, correspondant à des résumés de conférence, de présentation poster ou de chapitre de livre, ont été écartés.

Nous avons ensuite éliminé deux articles dont la revue avait un facteur d'impact inférieur à 1 pour l'un et non déterminé pour l'autre.

Après lecture intégrale, 11 articles ont été exclus car ils portaient sur des tests d'algorithmes utilisant uniquement des images provenant de bases de données. Ensuite, 3 autres articles ont été éliminés : l'un parce qu'il traitait du développement d'une IA pour identifier tous les grains de beauté présents sur des images dermatologiques téléchargées sur une plateforme de télémédecine, un autre qui cherchait à valider une méthode pour compter les naevus de manière standardisée avec seulement 10 sujets inclus et le dernier qui développait un logiciel pour détecter et quantifier des changements dans les lésions cutanées pigmentées (en particulier des naevus à partir de photographies numériques sérielles).

Un autre article a été éliminé car il validait une application mobile de santé pour le dépistage du cancer de la peau, suggérant une utilisation potentielle par les patients avant consultation. Deux autres études ont été exclues car elles se concentraient sur des lésions cutanées non cancéreuses (eczéma, psoriasis), réduisant le nombre à 28 articles. Deux études supplémentaires ont été retirées car elles visaient à déterminer l'épaisseur d'un mélanome ou son caractère invasif.

Enfin, une étude pilote portant sur le développement et l'évaluation clinique d'un outil de support à l'intelligence artificielle pour améliorer la qualité des photos en télémédecine a été exclue.

Les données concernant les articles exclus après lecture intégrale du texte sont disponibles dans l'annexe 2. Finalement, 25 articles ont été sélectionnés pour notre revue systématique. Le processus de sélection est illustré dans le diagramme de flux suivant :

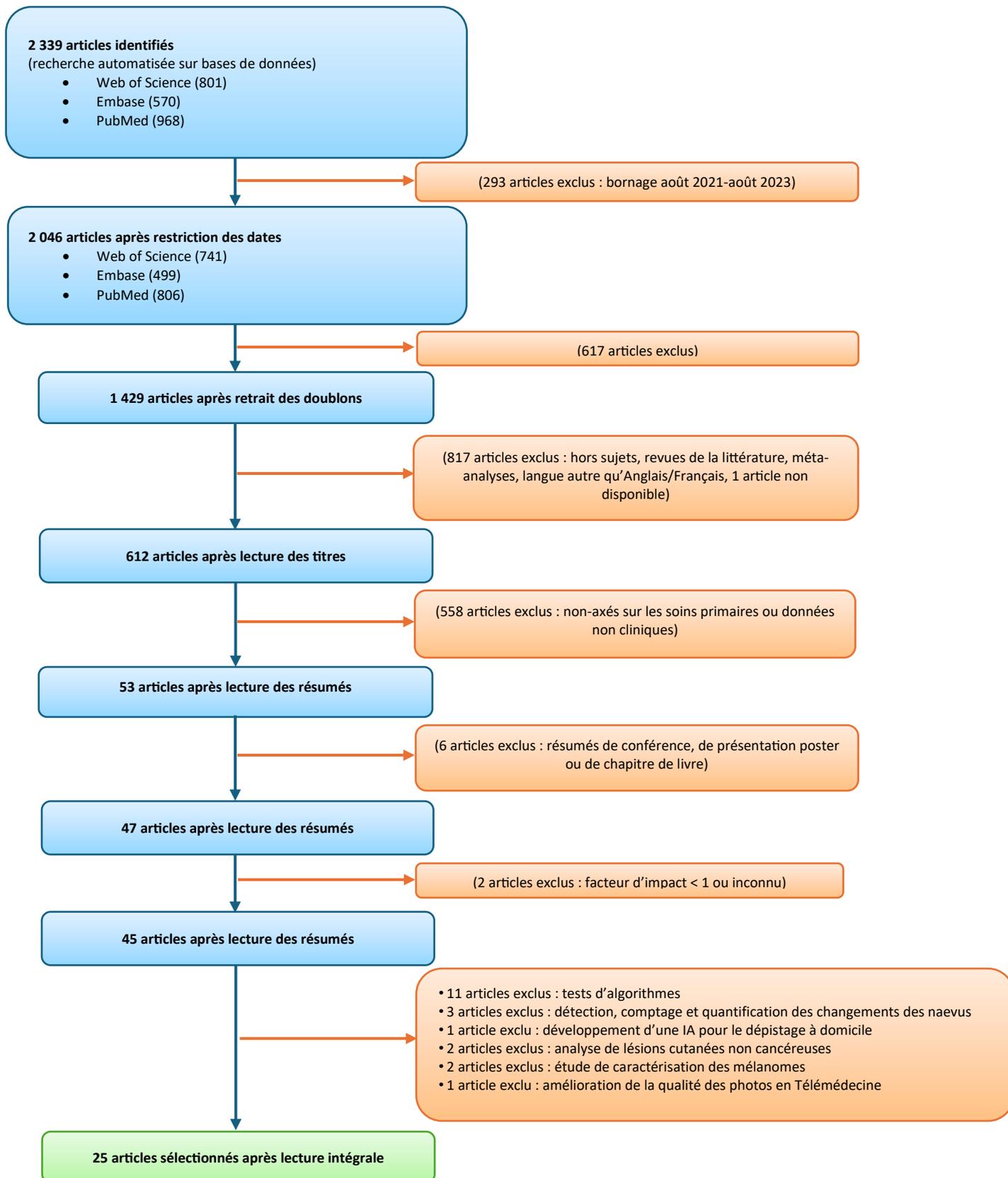


Figure 21 : Diagramme de flux de sélection des articles.

III - 2 ARTICLES INCLUS ET RÉSULTATS PRINCIPAUX :

Le tableau suivant résume les principales caractéristiques des articles sélectionnés (titres, auteurs, année de publication, etc...), leur évaluation, ainsi que les principaux résultats. Il contribue à répondre à notre question de recherche. Les articles sont classés par ordre alphabétique et par type d'étude (observationnelle, interventionnelle, randomisée ou pas).

Tableau 4 : Caractéristiques et évaluation des études sélectionnées

N°	Titre/Auteur/Année	Revue/Facteur d'impact	Pays	Méthodologie	Évaluation	Objectifs et Principaux résultats	Financement / Conflit d'intérêt
1	A deep learning based multimodal fusion model for skin lesion diagnosis using smartphone collected clinical images and metadata. Chubin Ou, Sitong Zhou, Ronghua Yang et al. 2022	Frontiers in Surgery / 1,8	Chine	Étude d'évaluation d'une méthode diagnostique / Observationnelle Rétrospective Transversale	STROBE 17/34 QUADAS-2 10/14	Objectif : Créer un outil accessible pour classifier les lésions cutanées à partir d'image et de métadonnées recueillies par Smartphone. Résultats : - L'inclusion des métadonnées (âge, type de peau, emplacement de la lésion...) a significativement amélioré les performances du modèle par rapport à l'utilisation des images seules (précision moyenne de 0,768 contre 0,616, AUROC de 0,947 contre 0,901). - Performances supérieures aux autres méthodes préexistantes d'intégration des métadonnées (exactitude équilibrée de 0,775 contre 0,765 pour MetaBlock, et 0,742 pour MetaNet).	National Natural Science Foundation of China, Guangdong Basic and Applied Basic Research Foundation, Medical Scientific Research Foundation of Guangdong Province / Aucun conflit d'intérêt déclaré
2	A reinforcement learning model for AI-based decision support in skin cancer. Catarina Barata, Veronica Rotemberg, Noel C. F. Codella et al. 2023	Nature Medicine / 82,9	Australie, Autriche, Espagne, États-Unis, Grèce, Italie, Portugal, Turquie	Étude d'évaluation d'une méthode diagnostique / Observationnelle Rétrospective et comparative de performance	STROBE 19/34 QUADAS-2 11/14	Objectif : Déterminer si les préférences humaines peuvent améliorer la précision des systèmes d'IA (attribution de poids influençant le modèle d'apprentissage par renforcement pour qu'il accorde une importance plus ou moins grande aux caractéristiques associées aux mélanomes ou aux lésions bénignes). Résultats : - Le modèle d'apprentissage par renforcement a montré une amélioration significative de la sensibilité pour le mélanome (de 61,4 % à 79,5 %) et pour le carcinome basocellulaire (de 79,4 % à 87,1 %). - A réduit la confiance excessive de l'IA dans les prédictions incorrectes tout en maintenant une précision élevée - A augmenté le taux de diagnostics corrects par les dermatologues de 12 % et amélioré les décisions de gestion optimale (de 57,4 % à 65,3 %).	FCT project and multiyear funding, MSK Cancer Center Support Grant / Silverchair, FotoFinder, Lilly, Novartis, MetaOptima Technology Inc, First Derm, Inhabit Brands Inc. Derma Medical Systems, Microsoft, Canfield Scientific Inc., Scibase AB, MoleMap NZ Limited, e-derm consult GmbH, Canfield Scientific Inc., MoleMap Australia Pty Ltd, Blaze Bioscience Inc.
3	Artificial Intelligence vs Medical Providers in the Dermoscopic Diagnosis of Melanoma. Jane M. Anderson, Izhaar Tejani, Tory Jarman et al. 2023	Cutis / 1,46	Canada, États-Unis	Étude d'évaluation d'une méthode diagnostique / Observationnelle Transversale et comparative de performance	STROBE 16/34 QUADAS-2 9/14	Objectif : Évaluer la performance d'une application d'IA dans la classification des lésions cutanées, en comparant ses résultats avec ceux obtenus par des dermatologues, des médecins généralistes et des prestataires de soins primaires de niveau intermédiaire Résultats : - L'application d'IA <i>Triage</i> a obtenu une sensibilité de 80 %, une spécificité de 95 % et une exactitude de 92 % dans la classification des lésions cutanées comme bénignes ou malignes, surpassant les performances des dermatologues (77 %, 57 % et 61 % respectivement) et autres professionnels de santé (78 %, 41 % et 48% respectivement pour les médecins généralistes) - L'IA a obtenu une valeur prédictive positive (PPV) de 80 % et une valeur prédictive négative (NPV) de 95 %, surpassant les dermatologues qui ont obtenu une PPV de 32 % et une NPV de 92 %, ainsi que les médecins généralistes avec une PPV de 25 % et une NPV de 90 %.	Aucun Financement spécifique extérieur / Deux auteurs sont employés par Triage Technologies Inc. (développeur de l'application d'IA évaluée dans cette étude)

N°	Titre/Auteur/Année	Revue/Facteur d'impact	Pays	Méthodologie	Évaluation	Objectifs et Principaux résultats	Financement / Conflit d'intérêt
4	<p>Assessing the generalisability of deep learning models trained on standardised and nonstandardised images and their performance against teledermatologists : retrospective comparative study.</p> <p>Ayooluwatomiwa I Oloruntoba, Tine Vestergaard, Toan D Nguyen et al. 2022</p>	<p>JMIR Dermatology / 3,2</p>	Australie	<p>Étude d'évaluation d'une méthode diagnostique / Observationnelle Rétrospective et comparative de performance</p>	<p>STROBE 17/34 QUADAS-2 11/14</p>	<p><u>Objectif</u> : Évaluer la généralisabilité des modèles de réseaux CNN entraînés sur des images (dermoscopiques et macroscopiques) de cancer de la peau standardisées (uniformité des paramètres de qualité d'image tels que la netteté, la luminosité, la polarisation, le grossissement, la couleur...) et non standardisées (dermoscopiques). <u>Résultats</u> :</p> <ul style="list-style-type: none"> - Les modèles entraînés sur des images standardisées ont surpassé le modèle non standardisé sur tous les ensembles de test (ex sur l'ensemble de test danois, le modèle CNN-S (standardisé) a atteint une AUROC de 0.861, tandis que le modèle non standardisé (CNN-NS) a obtenu une AUROC de 0.759). - L'un des CNN entraîné sur images standardisées, comparé aux télédermatologues, a montré une sensibilité et une spécificité comparables, bien qu'il ait été légèrement inférieur en performance générale. - Les modèles de CNN entraînés sur des images standardisées ont une meilleure généralisabilité (performance sur différents ensembles de données externes, y compris l'ensemble ISIC 2020, où le CNN-S a atteint une AUROC de 0.828 contre 0.763 pour le CNN-NS). 	<p>Australian Government Research Training Program Scholarship, NVIDIA Artificial Intelligence Fellowship, Monash-Airdoc Research Centre collaboration, National Health and Medical Research Council, Australian Cancer Research Foundation / MoleMap NZ Ltd, e-derm Consult GmbH, Canfield Scientific Inc., MoleMap Australia Pty Ltd, Blaze Bioscience Inc., First Derm, Novartis, Bristol Myers Squibb, Merck, Janssen, L'Oréal.</p>
5	<p>Assessment of Diagnostic Performance of Dermatologists Cooperating With a Convolutional Neural Network in a Prospective Clinical Study: Human With Machine.</p> <p>Julia K. Winkler, Andreas Blum, Katharina Kommos et al. 2023</p>	<p>JAMA Dermatology / 10,9</p>	Allemagne	<p>Étude d'évaluation d'une méthode diagnostique / Étude Observationnelle Prospective de Cohorte</p>	<p>STROBE 21/34 QUADAS-2 10/14</p>	<p><u>Objectif</u> : Évaluer si les dermatologues pouvaient améliorer leur performance diagnostique lorsqu'ils coopèrent avec un CNN pour la classification des lésions mélanocytaires <u>Résultats</u> :</p> <ul style="list-style-type: none"> - Les dermatologues seuls ont atteint une sensibilité moyenne de 84,2 % et une spécificité moyenne de 72,1 %. Lorsque les dermatologues ont utilisé les résultats du CNN, la sensibilité est passée à 100 % et la spécificité à 83,7 %. - L'utilisation du CNN a permis de réduire de 19,2 % le nombre d'excisions inutiles de névi bénins. 	<p>Aucun Financement spécifique extérieur / FotoFinder Systems, Amgen, Bristol Myers Squibb, MSD, Philochem, Roche, HEINE Optotechnik, SciBase, Magnosco</p>
6	<p>Clinical Utility of an AI-powered, Handheld Elastic Scattering Spectroscopy Device on the Diagnosis and Management of Skin Cancer by Primary Care Physicians.</p> <p>Erik Jaklitsch, Todd Thames, Thomaz de Campos Silva et al. 2023</p>	<p>Journal of Primary Care & Community Health / 3</p>	Australie, Etats-Unis	<p>Étude d'évaluation d'une méthode diagnostique / Observationnelle, Rétrospective et comparative de performance</p>	<p>STROBE 19/34 QUADAS-2 10/14</p>	<p><u>Objectifs</u> : Mesurer l'impact d'un dispositif portable utilisant la spectroscopie de diffusion élastique et l'IA sur la capacité des médecins généralistes à diagnostiquer et gérer (décision de biopsie ou de référence) les cancers de la peau. <u>Résultats</u> :</p> <ul style="list-style-type: none"> - L'utilisation du dispositif ESS a montré une augmentation significative de la sensibilité diagnostique (67 % vs 88 %) et de gestion (81 % vs 94 %) des lésions cutanées suspectes. - Les médecins ont montré une augmentation significative de leur confiance dans les décisions de gestion lorsqu'ils utilisaient le dispositif (p < 0,0001). - Le dispositif a permis de réduire le nombre de faux négatifs de 35 cas (de 42 à 7) pour le mélanome. 	<p>DermaSensor Inc. / DermaSensor, Inc., DermTech, Castle Biosciences, SkinAnalytics</p>

N°	Titre/Auteur/Année	Revue/Facteur d'impact	Pays	Méthodologie	Évaluation	Objectifs et Principaux résultats	Financement / Conflit d'intérêt
7	Convolutional neural network assistance significantly improves dermatologists' diagnosis of cutaneous tumours using clinical images. Wei Ba, Huan Wu, Wei W. Chen et al. 2022	European Journal of Cancer / 8,4	Chine	Étude d'évaluation d'une méthode diagnostique / Observationnelle, Rétrospective et comparative de performance	STROBE 20/34 QUADAS-2 9/14	Objectif : Évaluer l'impact de l'assistance d'un CNN sur la performance des dermatologues dans l'identification des tumeurs cutanées à partir d'images cliniques. Résultats : - La précision des dermatologues est passée de 62,78 % sans assistance à 76,60 % avec l'assistance du CNN, avec une augmentation du kappa (concordance) de 0,59 à 0,74. - La sensibilité et la spécificité ont également augmenté avec l'assistance du CNN (respectivement de 83,21 % à 89,56 % et de 80,92 % à 87,90 %). - Les dermatologues moins expérimentés ont particulièrement bénéficié de l'assistance du CNN.	Aucun Financement spécifique extérieur / Aucun conflit d'intérêt déclaré
8	Diagnostic performance of augmented intelligence with 2D and 3D total body photography and convolutional neural networks in a high-risk population for melanoma under real-world conditions: A new era of skin cancer screening? Sara E. Cerminara, Phil Cheng, Lisa Kostner et al. 2023	European Journal of Cancer / 8,4	Suisse	Étude d'évaluation d'une méthode diagnostique / Observationnelle Prospective de Cohorte	STROBE 24/34 QUADAS-2 10/14	Objectif : Évaluer la performance des CNN 2D et 3D pour la classification des lésions mélanocytaires et examiner le potentiel de collaboration entre les cliniciens et l'IA dans la détection du mélanome en pratique clinique quotidienne. Résultats : - Les CNN 3D ont montré une sensibilité supérieure aux CNN 2D (90 % contre 70 %) et des performances comparables à celles des dermatologues, bien que ces derniers soient meilleurs en spécificité (92.3 % pour les dermatologues contre 64.6 % pour les CNN 3D). - Les CNN détectent efficacement les mélanomes, mais rencontrent des difficultés à classifier correctement les lésions suspectes (pas clairement bénigne ou maligne). - L'IA peut améliorer la précision des diagnostics des cliniciens, mais les erreurs potentielles des algorithmes peuvent également induire en erreur, y compris les plus expérimentés (spécificité diminuée à 86,15 % lorsque les dermatologues utilisent l'IA).	Département de Dermatologie des Hôpitaux Universitaires de Bâle et Zurich, Research Foundation for Young Researchers et Voluntary Academic Society de l'Université de Bâle / Aucun conflit d'intérêt déclaré
9	Early melanoma diagnosis with sequential dermoscopic images. Zhen Yu, Jennifer Nguyen, Toan D. Nguyen et al. 2022	IEEE Transactions on Medical Imaging / 10,6	Australie, Chine	Étude d'évaluation d'une méthode diagnostique / Observationnelle Rétrospective Transversale et comparative de performance	STROBE 17/34 QUADAS-2 9/14	Objectif : Créer un modèle automatisé qui utilise des images dermoscopiques séquentielles pour améliorer le diagnostic précoce du mélanome, tout en réduisant les biopsies inutiles, en prenant en compte les changements temporels et morphologiques des lésions cutanées. Résultats : - Le modèle informatique a surpassé les cliniciens expérimentés en termes de spécificité diagnostique (66,67 % contre 46,76 %), réduisant ainsi les excisions inutiles de lésions bénignes. - Détection plus précoce des mélanomes. - Bien que globalement performant, le modèle a été moins efficace que les cliniciens dans la détection des mélanomes invasifs, avec une sensibilité de 61,8 % contre 67,9 %.	Australian National Health and Medical Research Council, National Natural Science Foundation of China, Outstanding Young Scholars Support Program of China, Xi'an Jiaotong University, NHMRC Early Career Fellowship Grant / Aucun conflit d'intérêt déclaré

N°	Titre/Auteur/Année	Revue/Facteur d'impact	Pays	Méthodologie	Évaluation	Objectifs et Principaux résultats	Financement / Conflit d'intérêt
10	<p>Effectiveness of an imageanalysing artificial intelligence-based digital health technology to diagnose nonmelanoma skin cancer and benign skin lesions.</p> <p>Helen Marsden, Caroline Morgan , Stephanie Austin et al. 2023</p>	Frontiers in Medicine / 5,09	Royaume-Uni	Étude Observationnelle Prospective Transversale en aveugle, Multicentrique	STROBE 20/34 QUADAS-2 12/14	<p><u>Objectif</u> : Valider cliniquement la capacité d'un dispositif médical basé sur l'IA (AlaMD) à identifier les cancers de la peau non-mélanome, notamment le carcinome basocellulaire et le carcinome épidermoïde.</p> <p><u>Résultats</u> :</p> <ul style="list-style-type: none"> - AlaMD a montré une performance comparable à celle des dermatologues, avec des AUROC supérieurs à 0,85 pour les BCC et les SCC. - L'IA a atteint une sensibilité de plus de 90 % pour identifier les lésions malignes avec des images prises par différents smartphones. 	<p>Skin Analytics Ltd., Londres / Certains auteurs sont des employés de Skin Analytics et ont reçu des actions ou des options d'achat d'actions de l'entreprise</p>
11	<p>Exploring the potential of artificial intelligence in improving skin lesion diagnosis in primary care.</p> <p>Anna Escalé-Besa, Oriol Yéla-mos, Josep Vidal-Alaball et al. 2023</p>	Scientific Reports / 4,6	Espagne, Etats-Unis, Suède	Étude d'évaluation d'une méthode diagnostique / Observationnelle Prospective de Cohorte, Multicentrique	STROBE 23/34 QUADAS-2 11/14	<p><u>Objectifs</u> : Tester la faisabilité d'un modèle d'apprentissage automatique pour identifier 44 maladies cutanées dans un cadre de soins primaires/Comparer la précision du modèle avec celle des médecins généralistes et de la télédermatologie.</p> <p><u>Résultats</u> :</p> <ul style="list-style-type: none"> - La précision diagnostique du modèle de Machine Learning (ML) pour les maladies de la peau était inférieure à celle des médecins généralistes et des dermatologues, avec une précision Top-1 de 39 % pour le ML contre 64 % pour les médecins généralistes et 72 % pour les dermatologues. - Lorsque le modèle ML était limité aux diagnostics pour lesquels il avait été formé, sa précision Top-1 augmentait à 48% et sa précision Top-5 atteignait 89% (comparable à celle des dermatologues). - 92% des médecins généralistes ont trouvé l'outil ML utile comme support au diagnostic différentiel. 	<p>Ajut d'alliberació (libération grant), Société Espagnole de Médecine de Famille et Communautaire (semFYC), Fundació Institut de Recerca i Innovació en Ciències de la Vida i de la Salut a la Catalunya Central (IRIS-CC) et Consultori Bayés Vic, S.L. / L'un des auteurs est fondateur de First Derm (proposant Skin Image Search app gratuitement), un autre en est un employé.</p>
12	<p>Implementation of artificial intelligence algorithms for melanoma screening in a primary care setting.</p> <p>Mara Giavina-Bianchi, Raquel Machado de Sousa, Vitor Zago de Almeida Paciello et al. 2021</p>	PLOS ONE / 3,8	Brésil	Étude d'évaluation d'une méthode diagnostique / Observationnelle Rétrospective Transversale et comparative de performance	STROBE 17/34 QUADAS-2 9/14	<p><u>Objectifs</u> : Créer un système de diagnostic assisté par ordinateur (CAD) capable d'analyser à la fois des images cliniques et dermoscopiques, en fournissant des recommandations de gestion des lésions aux médecins généralistes.</p> <p><u>Résultats</u> :</p> <ul style="list-style-type: none"> - Les algorithmes ont montré de bonnes performances, avec une précision de 84 à 89 % pour les images cliniques et dermoscopiques respectivement. - La sensibilité des algorithmes était proche de 90 %, et la spécificité était de 83 % pour les images cliniques et 89 % pour les images dermoscopiques. - L'étude inclut des retours d'expérience des médecins généralistes, qui ont généralement accueilli favorablement l'outil de diagnostic assisté par ordinateur. 	<p>Ministère de la Santé du Brésil / Aucun conflit d'intérêt déclaré</p>

N°	Titre/Auteur/Année	Revue/Facteur d'impact	Pays	Méthodologie	Évaluation	Objectifs et Principaux résultats	Financement / Conflit d'intérêt
13	Implementation of artificial intelligence for the detection of cutaneous melanoma within a primary care setting: prevalence and types of skin cancer in outdoor enthusiasts. Ian J. Miller, Michael Stapelberg, Nedeljka Rosic et al. 2023	PeerJ / 3,06	Australie	Étude d'évaluation d'une méthode diagnostique / Observationnelle Prospective Transversale	STROBE 21/34 QUADAS-2 13/14	<u>Objectifs</u> : - Évaluer l'incidence des cancers de la peau chez des individus pratiquant des activités en extérieur et comparer leur prévalence avec celle d'autres populations étudiées dans la littérature. - Analyser l'efficacité d'un dispositif d'imagerie non invasif basé sur l'IA (FotoFinder) pour détecter le mélanome et d'autres cancers de la peau. <u>Résultats</u> : - Les taux de mélanome (OR 119,8) et d'autres cancers cutanés (BCC et SCC) étaient significativement plus élevés chez les surfeurs et les marcheurs/coueurs comparés à la population générale australienne et à d'autres groupes similaires (exposés au soleil) tels que les skieurs et golfeurs. - Les résultats pour la sensibilité et la spécificité de l'IA (FotoFinder) étaient inférieurs aux attentes, avec une sensibilité de 53,5 % et une spécificité de 54,4 %, comparativement à des études antérieures rapportant des valeurs plus élevées (sensibilité moyenne de 84,2 %, spécificité moyenne de 89,1 %).	Johnson and Johnson pour l'achat du dermatoscope numérique haute résolution. / Un des auteurs est éditeur de section pour PeerJ (Sports Medicine and Rehabilitation)
14	Lesion identification and malignancy prediction from clinical dermatological images. Meng Xia, Meenal K. Kheterpa, Samantha C. Wong et al. 2022	Scientific Reports / 4,6	États-Unis	Étude d'évaluation d'une méthode diagnostique / Observationnelle Rétrospective et comparative de performance	STROBE 15/34 QUADAS-2 10/14	<u>Objectifs</u> : Développer et évaluer un système d'aide à la décision clinique basé sur une IA capable d'identifier et de classer des lésions cutanées sur des images de vue d'ensemble. <u>Résultats</u> : - Pour les images de dermoscopie, le modèle a obtenu une AUROC de 0,779, montrant une performance comparable à celle de deux des trois dermatologues, avec une légère infériorité par rapport au dermatologue le plus expérimenté. - Pour les images obtenues à partir de smartphones, le modèle a atteint une AUROC de 0,789, se plaçant également à un niveau similaire ou supérieur à deux des trois dermatologues, indiquant sa robustesse même avec des images de qualité variable.	Aucun financement spécifique extérieur / Aucun conflit d'intérêt déclaré
15	MobileSkin: Classification of Skin Lesion Images Acquired Using Mobile Phone-Attached Hand-Held Dermoscopes. Abdurrahim Yilmaz, Gulsum Gencoglan, Rahmetullah Varol et al. 2022	Diagnostics / 3,9	Turquie	Étude d'évaluation d'une méthode diagnostique / Observationnelle Rétrospective et comparative de performance	STROBE 14/34 QUADAS-2 9/14	<u>Objectifs</u> : Développer et évaluer des modèles d'apprentissage profond capables de classer automatiquement des images dermoscopiques (prises à l'aide d'un dermatoscope portable attaché à un téléphone mobile) de lésions cutanées. <u>Résultats</u> : - Les modèles d'apprentissage profond les plus performants étaient <i>MobileNetV2</i> et <i>Xception</i> , avec des précisions de classification respectives de 89,18 % et 89,64 %. - Les résultats indiquent que ces modèles peuvent classer les lésions cutanées avec une précision élevée, proche de celle des dermatologues expérimentés.	Aucun financement spécifique extérieur / Aucun conflit d'intérêt déclaré

N°	Titre/Auteur/Année	Revue/Facteur d'impact	Pays	Méthodologie	Évaluation	Objectifs et Principaux résultats	Financement / Conflit d'intérêt
16	Monitoring patients at risk for melanoma: May convolutional neural networks replace the strategy of sequential digital dermoscopy? Julia K. Winkler, Philipp Tschandl, Ferdinand Toberer et al. 2021	European Journal of Cancer / 8,4	Allemagne Autriche	Étude d'évaluation d'une méthode diagnostique / Observationnelle Rétrospective Transversale et comparative de performance	STROBE 15/34 QUADAS-2 10/14	<u>Objectif</u> : Examiner si les CNN peuvent remplacer la stratégie de la dermoscopie numérique séquentielle (SDD) pour la détection précoce du mélanome. <u>Résultats</u> : - Les CNN ont montré une capacité limitée à identifier les mélanomes sur les images de base, avec des taux de classification corrects de 15,3 % pour CNN-1 et 13,6 % pour CNN-2, bien que leur performance se soit améliorée après le suivi de la SDD, atteignant des sensibilités de 44,1 % pour CNN-1 et 49,2 % pour CNN-2. - Les dermatologues ont surpassé les CNN dans la classification correcte des ensembles d'images de base avec une moyenne de 40,7 %, mais une approche collaborative entre les deux (en prenant en compte les données de suivi SDD) a permis d'augmenter la sensibilité diagnostique à 78,2 %, bien que la spécificité ait baissé à 54,2 %.	Aucun financement spécifique extérieur / Aucun conflit d'intérêt déclaré mais liens avec MetaOptima, Lilly, Silverchair, FotoFinder, Biotest, MSD, Janssen-Cilag, BMS, 3Gen, Derma Medical, Scibase AB, Systems GmbH, Heine Optotechnik, Magnosco GmbH
17	New AI-algorithms on smartphones to detect skin cancer in a clinical setting-A validation study. Teresa Kränke, Katharina Tripolt-Droschl, Lukas Röd et al. 2023	PLOS ONE / 3,8	Autriche	Étude d'évaluation d'une méthode diagnostique / Observationnelle Prospective de Cohorte et comparative de performance	STROBE 18/34 QUADAS-2 12/14	<u>Objectif</u> : Évaluer la précision diagnostique et la capacité d'évaluation des risques de deux nouveaux CNN (<i>analyze</i> et <i>detect</i>) intégrés dans des smartphones, afin de détecter le potentiel malin de diverses lésions cutanées dans un cadre clinique. <u>Résultats</u> : - L'algorithme <i>detect</i> a montré une précision diagnostique globale de 88 %, surpassant l'algorithme <i>analyze</i> (82 %), avec une sensibilité supérieure à 95 % pour les deux algorithmes. - La spécificité de <i>detect</i> était également supérieure à celle de <i>analyze</i> (94,85 % contre 90,32 %)	Amiflow Ltd / Aucun conflit d'intérêt déclaré
18	Over-Detection of Melanoma suspect Lesions by a CE-Certified Smartphone App: Performance in Comparison to Dermatologists, 2D and 3D Convolutional Neural Networks in a Prospective Data Set of 1204 Pigmented Skin Lesions Involving Patients' Perception. Anna Sophie Jahn, Alexander Andreas Navarini, Sara Elisa Cerminara et al. 2022	Cancers / 5,2	Suisse	Étude d'évaluation d'une méthode diagnostique / Observationnelle Prospective de Cohorte et comparative de performance	STROBE 20/34 QUADAS-2 11/14	<u>Objectif</u> : Valider prospectivement l'efficacité diagnostique de l'application pour smartphone SkinVision®, utilisant un algorithme d'IA pour la stratification du risque de mélanome. <u>Résultats</u> : - L'application SkinVision® a montré une sensibilité variant de 41,3 % à 83,3 %, une spécificité de 60 % à 82,9 % et une aire sous la courbe AUC de 0,62 à 0,72, en fonction du standard de référence utilisé (combinaison d'évaluations des dermatologues et d'IA versus histopathologie). - Comparée aux dermatologues, qui ont démontré une spécificité plus élevée de 92,7 % et une sensibilité variant selon leur niveau d'expérience (jusqu'à 100 % pour les plus expérimentés), l'application présente un taux élevé de faux positifs.	Fonds institutionnels des départements de dermatologie des Hôpitaux Universitaires de Bâle et de Zurich / Aucun conflit d'intérêt déclaré

N°	Titre/Auteur/Année	Revue/Facteur d'impact	Pays	Méthodologie	Évaluation	Objectifs et Principaux résultats	Financement / Conflit d'intérêt
19	<p>Prospective validation of dermoscopy-based opensource artificial intelligence for melanoma diagnosis (PROVE-AI study).</p> <p>Michael A. Marchetti, Emily A. Cowen, Nicholas R. Kurtansky et al. 2023</p>	NPJ Digital Medicine / 15,2	États-Unis	Étude d'évaluation d'une méthode diagnostique / Observationnelle Prospective de Cohorte et comparative de performance	STROBE 21/34 QUADAS-2 13/14	<p><u>Objectifs</u> :</p> <ul style="list-style-type: none"> - Valider prospectivement l'exactitude d'un algorithme d'IA open-source, appelé ADAE, pour le diagnostic du mélanome à partir d'images dermoscopiques de lésions cutanées. - Évaluer l'impact potentiel de l'algorithme sur la prise de décision des dermatologues, notamment en termes de réduction des biopsies inutiles tout en maintenant une haute sensibilité pour détecter les mélanomes. <p><u>Résultats</u> :</p> <ul style="list-style-type: none"> - L'algorithme ADAE a montré une sensibilité de 96,8 % et une spécificité de 37,4 %, surpassant les dermatologues dans la discrimination entre mélanomes et lésions bénignes tout en réduisant le nombre de biopsies inutiles. - Après l'exposition aux résultats de l'algorithme, les décisions de gestion des dermatologues ont permis de ne pas biopsier environ 33,7 % des lésions non-mélanomes. - Première étude à valider un algorithme d'IA open-source pour le diagnostic du mélanome dans un cadre clinique réel, avec une prise de décision immédiate. 	<p>National Institutes of Health, National Cancer Institute Cancer Center Support Grant, Melanoma Research Alliance Young Investigator Award, Harry J. Lloyd Charitable Trust Melanoma Research Grant, National Center for Advancing Translational Sciences grant /</p> <p>Aucun conflit d'intérêt déclaré mais liens avec Inhabit Brands, Canfield Scientific Inc., Janssen Research and Development, HCW Health LLC et SKIP Derm LLC.</p>
20	<p>Superiority of artificial intelligence compared to dermatologists and primary care providers in the diagnosis of malignant melanoma.</p> <p>Jane M. Anderson, Izhaar Tejani, Tory Jarmain et al. 2022</p>	IEEE Transactions On Medical Imaging / 10,6	Canada, États-Unis	Étude d'évaluation d'une méthode diagnostique / Observationnelle Rétrospective Transversale et de performance comparative	STROBE 16/34 QUADAS-2 9/14	<p><u>Objectif</u> : Évaluer la performance de l'application d'IA <i>Triage</i> pour classer les lésions cutanées comme bénignes ou malignes et comparer cette performance à celle des dermatologues et des prestataires de soins primaires.</p> <p><u>Résultats</u> :</p> <ul style="list-style-type: none"> - L'application <i>Triage</i> a surpassé les professionnels de la santé dans la spécificité (95 %), la précision (92 %) et la valeur prédictive positive (80 %), tout en maintenant une sensibilité comparable à celle des dermatologues. - L'IA a montré une meilleure sensibilité par rapport aux prestataires de soins de niveau intermédiaire (infirmiers praticiens et assistants médicaux). 	<p>Aucun financement spécifique extérieur /</p> <p>Aucun conflit d'intérêt déclaré</p>

N°	Titre/Auteur/Année	Revue/Facteur d'impact	Pays	Méthodologie	Évaluation	Objectifs et Principaux résultats	Financement / Conflit d'intérêt
21	<p>Validation of artificial intelligence prediction models for skin cancer diagnosis using dermoscopy images: the 2019 International Skin Imaging Collaboration Grand Challenge. Marc Combalia, Noel Codella, Veronica Rotemberg et al. 2022</p>	<p>The Lancet Digital Health / 30,8</p>	<p>Autriche Espagne États-Unis</p>	<p>Etude d'évaluation d'une méthode diagnostique / Observationnelle Rétrospective Transversale et de performance comparative</p>	<p>STROBE 20/34 QUADAS-2 8/14</p>	<p><u>Objectif</u> : Valider les performances des modèles de prédiction basés sur l'IA pour le diagnostic du cancer de la peau en utilisant des images dermoscopiques. <u>Résultats</u> :</p> <ul style="list-style-type: none"> - L'algorithme le plus performant a atteint une précision équilibrée de 63,6 % sur l'ensemble de test, en baisse par rapport aux 86,1 % rapportés sur l'ensemble de données HAM10000. - Concernant les catégories non entraînées (NT, catégories de maladies ou de lésions qui n'étaient pas incluses dans l'ensemble de données utilisées pour former les algorithmes), 89 % des cas ont été mal classés, avec 47 % des images NT incorrectement identifiées comme des diagnostics malins, ce qui pourrait entraîner des biopsies inutiles. - Les algorithmes ont surpassé les dermatologues dans plusieurs catégories diagnostiques, notamment pour les carcinomes basocellulaires (91 % de précision vs 70 %), les dermatofibromes (73 % vs 50 %) et les nævus (76 % vs 56 %), mais les dermatologues ont mieux réussi pour la catégorie NT, avec 26 % de classifications correctes contre 6 % pour les algorithmes. 	<p>Melanoma Research Alliance, La Marató de TV3 / Aucun conflit d'intérêt déclaré</p>
22	<p>Artificial intelligence in mobile health for skin cancer diagnostics at home (AIM HIGH): a pilot feasibility study. Anna M. Smak Gregoor, Tobias E. Sangers, Just AH. Eekhof et al. 2023</p>	<p>eClinical Medicine / 15,1</p>	<p>Pays-Bas</p>	<p>Etude Interventionnelle de Cohorte Non Randomisée / Étude de faisabilité pilote</p>	<p>TREND 35/58 MINORS 11/16</p>	<p><u>Objectif</u> : Évaluer la faisabilité de l'intégration d'une application mobile basée sur l'IA pour le diagnostic du cancer de la peau dans les soins primaires, en examinant les conditions nécessaires pour mener une telle étude et en évaluant l'impact potentiel de cette technologie sur le parcours de soins des patients. <u>Résultats</u> :</p> <ul style="list-style-type: none"> - L'application mHealth pour le diagnostic du cancer de la peau a correctement identifié 80 % des lésions bénignes et 90,9 % des lésions (pré)malignes, suggérant une bonne performance diagnostique. - 54 % des patients ayant reçu une évaluation de faible risque par l'application ont déclaré qu'ils auraient annulé leur visite chez le généraliste. - Bien que les généralistes n'aient pas modifié leur diagnostic initial, l'application a conduit à un ajustement du plan de traitement dans 10 % des cas. 	<p>SkinVision B.V / Le département de dermatologie a reçu un financement de recherche non restrictif de SkinVision B.V, un auteur déclaré des liens avec Pfizer, Janssen-Cilag et UCB</p>

N°	Titre/Auteur/Année	Revue/Facteur d'impact	Pays	Méthodologie	Évaluation	Objectifs et Principaux résultats	Financement / Conflit d'intérêt
23	Augmenting the accuracy of trainee doctors in diagnosing skin lesions suspected of skin neoplasms in a realworld setting: A prospective controlled beforeandafter study. Young Jae Kim, Jung-Im Na, Seung Seog Han et al. 2022	PLOS ONE / 3,8	Corée du Sud	Étude d'évaluation d'une méthode diagnostique / Étude pré-post Interventionnelle Contrôlée	TREND 33/58 MINORS 10/16	Objectif : Évaluer si un algorithme d'IA peut améliorer la précision des médecins en formation dans le diagnostic des lésions cutanées suspectées de néoplasmes cutanés dans un contexte réel. Résultats : - L'assistance par IA a amélioré la précision du diagnostic des médecins en formation, avec une précision Top-1 (le diagnostic principal est correct) passant de 46,5 % à 58,3 % dans le groupe IA. - Le nombre de diagnostics différentiels a également augmenté de 1,9 à 2,2 par cas après l'utilisation de l'IA. - Le groupe contrôle n'a pas montré d'amélioration significative de la précision après une simple révision des photographies, soulignant l'efficacité spécifique de l'algorithme d'IA.	Aucun Financement spécifique extérieur / Un des auteurs a fondé IDerma, Inc.
24	Improving skin cancer management with ARTificial intelligence: A pre-post intervention trial of an artificial intelligence system used as a diagnostic aid for skin cancer management in a real-world specialist dermatology setting. Claire Felmingham, Samantha MacNamara, William Cranwell et al. 2021	Journal of the American Academy of Dermatology / 13,8	Australie	Étude pré-post Interventionnelle Non Randomisée, Monocentrique / Étude comparative de performance	TREND 34/58 MINORS 21/24	Objectif : Évaluer l'efficacité d'un CNN développé pour aider au diagnostic du cancer de la peau en comparant ses performances avec celles des télédermatologues dans un cadre réel, afin de déterminer si son utilisation pouvait réduire la charge de travail des dermatologues et en améliorer l'accès. Résultats : - Le CNN présentait une aire sous la courbe AUROC (performance diagnostique) de 0,837, ce qui est légèrement supérieur à celle des télédermatologues (0,807). - Le CNN a enregistré un taux de faux négatifs de 3,7 %, ce qui est inférieur au taux de 9,3 % des télédermatologues, mais supérieur au taux de 0,6 % des dermatologues traitants.	Victorian Medical Research Acceleration Fund, MoleMap Ltd / Castle Biosciences, MoleMap Ltd, Proctor & Gamble, L'Oreal.
25	Evaluation of Artificial Intelligence-Assisted Diagnosis of Skin Neoplasms: A SingleCenter, Paralleled, Unmasked, Randomized Controlled Trial. Seung Seog Han, Young Jae Kim, Ik Jun Moon et al. 2022	Journal of Investigative Dermatology / 6,5	Corée du Sud	Etude d'évaluation d'une méthode diagnostique / Interventionnelle Randomisée Monocentrique	CONSORT 26/37 Jadad 3/5	Objectif : Évaluer l'efficacité d'un algorithme d'IA pour améliorer la précision diagnostique des médecins non spécialisés en dermatologie. L'étude visait à déterminer si l'IA pouvait assister ces professionnels de santé en conditions réelles, particulièrement en milieu de soins primaires, et à comparer les performances de l'IA avec celles des dermatologues confirmés. Résultats : - L'IA a significativement amélioré la précision diagnostique des médecins non spécialistes en dermatologie, avec une précision passant de 30,7 % à 54,7 % ($p < 0,0001$) pour les diagnostics principaux. - L'amélioration de précision chez les résidents en dermatologie n'était pas significative, passant de 51,7 % à 53,1 % ($p = 0,86$). - La précision de l'IA combinée aux médecins non spécialistes était comparable à celle des dermatologues atteignant une précision de 56 % (pour les cas confirmés par biopsie).	Aucun Financement spécifique extérieur / Un des auteurs a fondé IDerma, Inc.

Dans l'annexe 3, nous présentons l'évaluation de la qualité de chaque rapport de recherche, réalisée grâce aux échelles STROBE (Strengthening The Reporting of Observational Studies in Epidemiology), TREND (Transparent Reporting of Evaluations with Nonrandomized Designs) et Consolidated Standards of Reporting Trials (CONSORT), les données étant présentées de façon exhaustive.

Dans l'annexe 4, la qualité méthodologique de chaque étude est évaluée de façon détaillée dans les échelles QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies) (58), de MINORS (Methodological index for non-randomized studies) (59) et du score de Jadad (58).

Le « CanTest Framework » établit les phases de développement nécessaires pour garantir que les nouveaux tests de diagnostic ou les nouvelles technologies soient adaptés à leur(s) objectif(s) lorsqu'ils sont introduits dans la pratique clinique (62). Il vise à éviter leur implémentation prématurée qui pourrait entraîner un dysfonctionnement dans des populations à faible prévalence (comme les soins primaires). Le CanTest Framework souligne également l'importance de comprendre comment un test s'intègre dans des stratégies diagnostiques plus larges et influence les décisions thérapeutiques ainsi que les résultats des patients.

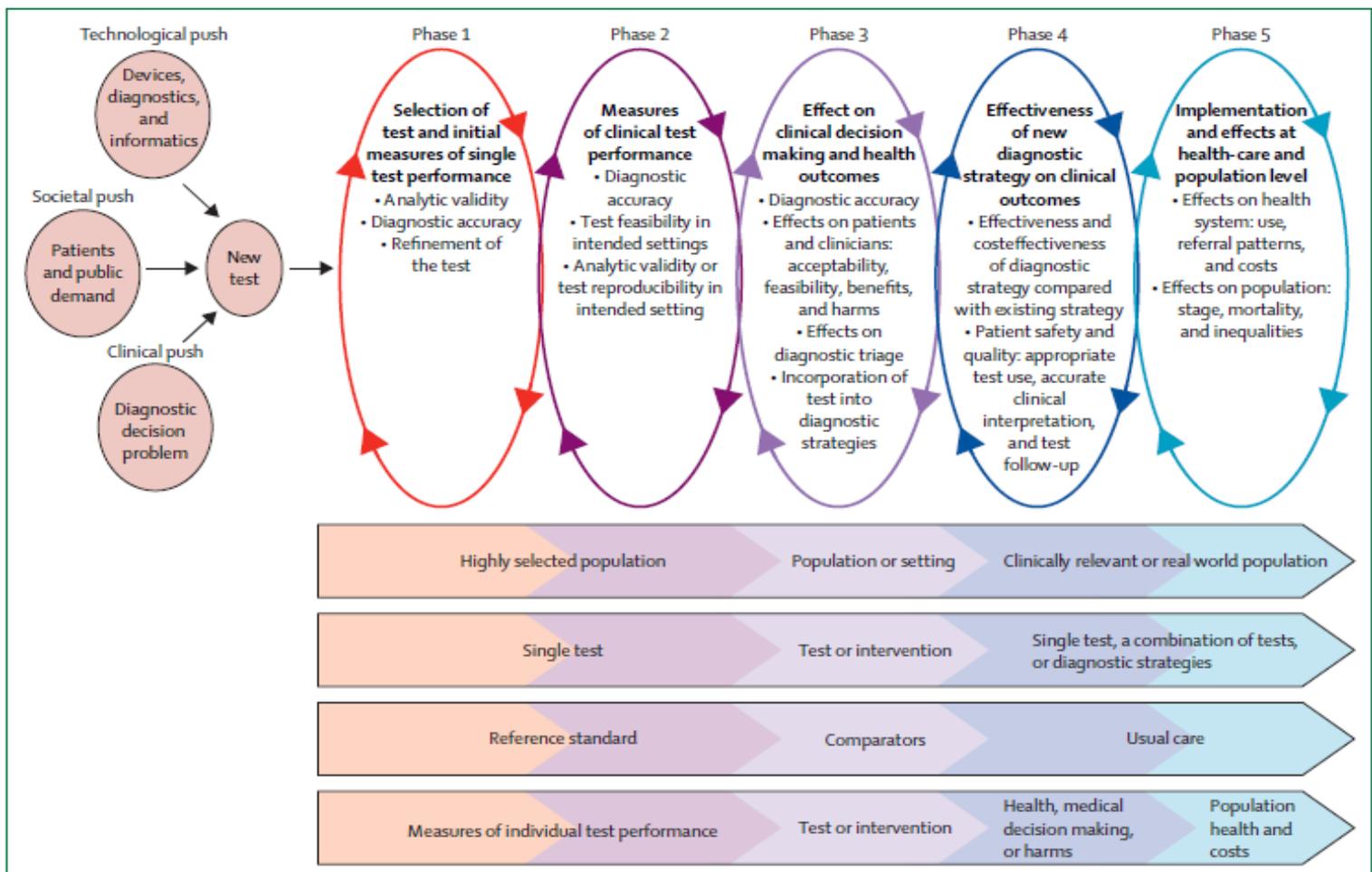


Figure 22 : Le CanTest Framework. D'après Jones et al : les phases de développement des tests de diagnostics sont représentées de gauche à droite.

Il s'agit d'un processus en plusieurs phases, allant de la création initiale du test jusqu'à son déploiement dans les systèmes de santé.

- Phase 1 : Sélection du test et premières mesures de performance.

Cette phase se concentre sur la *validité analytique* et l'*exactitude diagnostique* du nouveau test. L'objectif est d'affiner celui-ci en évaluant sa performance dans une population hautement sélectionnée, souvent dans des études de validation de principe. Cette étape vise à garantir qu'il fonctionne de manière fiable dans des conditions contrôlées.

- Phase 2 : Mesures de performance clinique du test.

La deuxième phase évalue la *faisabilité* du test dans des contextes cliniques réels. Cela inclut l'exactitude diagnostique et sa reproductibilité. La population étudiée ici se rapproche de celle dans laquelle il sera utilisé même si elle reste encore relativement sélectionnée.

- Phase 3 : Impact sur la prise de décision clinique et les résultats de santé.

L'accent est mis sur l'évaluation de l'impact du test sur les *décisions cliniques* et les *résultats pour les patients*. Cette phase examine comment il influence la prise en charge des patients, notamment la gestion du tri diagnostique et son incorporation dans des stratégies diagnostiques plus larges.

- Phase 4 : Efficacité de la nouvelle stratégie diagnostique sur les résultats cliniques.

Cette phase évalue l'*efficacité* de l'utilisation du nouveau test en tant que partie d'une stratégie diagnostique, comparée aux soins habituels. Elle examine également le rapport *coût-efficacité*, la sécurité des patients et la manière dont il s'intègre dans le système de soins standard.

- Phase 5 : Mise en œuvre et effets au niveau du système de santé et de la population.

La dernière phase se concentre sur les *effets à l'échelle du système de santé* de la mise en œuvre du test. Cela inclut l'impact sur les coûts, les schémas de recours aux soins et les résultats de santé de la population, telles que la mortalité et les inégalités en matière de santé.

Parmi les 25 articles analysés, 60 % appartiennent à la Phase 2 du CanTest Framework, qui évalue la performance clinique et la faisabilité des tests diagnostiques dans des environnements réels. Ces études se concentrent principalement sur la manière dont l'intelligence artificielle (IA) peut être utilisée pour améliorer le diagnostic des lésions cutanées en conditions cliniques, en évaluant des indicateurs tels que la sensibilité, la spécificité et la précision dans un cadre de test. La Phase 2 est souvent considérée comme une étape où les outils sont encore en phase de validation, testés dans des contextes simulés avant d'être intégrés plus largement dans la pratique clinique.

40 % des articles appartiennent à la Phase 3 qui évalue l'impact de la collaboration entre les médecins et l'IA sur les décisions cliniques et la gestion des patients, en particulier dans des environnements

réels. Ces études sont essentielles pour comprendre l'effet de l'IA sur la pratique quotidienne, notamment sur la réduction des erreurs diagnostiques et l'amélioration de la confiance des cliniciens.

À ce jour, nous n'avons trouvé aucun article concernant les Phases 4 et 5 du CanTest Framework dans le domaine de l'IA appliquée à l'analyse des lésions cutanées. Celles-ci portent sur l'efficacité à grande échelle et l'impact au niveau systémique de l'IA dans le domaine clinique, impliquant une intégration complète de ces outils dans les stratégies de santé publique et leur évaluation au niveau populationnel.

III – 3 ANALYSE DES RÉSULTATS

Pour réaliser une analyse des résultats des recherches sélectionnées la plus claire possible, nous avons regroupé les études en six thèmes, résumés dans le tableau ci-dessous :

Tableau 5 : Classement des articles sélectionnés par thème

CATÉGORIE D'ARTICLE	N°	TITRE DE L'ARTICLE	AUTEURS	ANNÉE
1. Analyse des lésions cutanées grâce à l'IA intégrée dans des applications pour Smartphones	1	A Deep Learning based multimodal fusion model for skin lesion diagnosis using smartphone collected clinical images and metadata.	Chubin Ou, Sitong Zhou, Ronghua Yang et al.	2022
	22	Artificial intelligence in mobile health for skin cancer diagnostics at home (AIM HIGH): a pilot feasibility study.	Anna M. Smak Gregoor, Tobias E. Sangers, Just AH. Eekhof et al.	2023
	17	New AI-algorithms on smartphones to detect skin cancer in a clinical setting-A validation study.	Teresa Kränke, Katharina Tripolt-Droschl, Lukas Röd et al.	2023
	18	Over-Detection of Melanoma-Suspect Lesions by a CE-Certified Smartphone App: Performance in Comparison to Dermatologists, 2D and 3D Convolutional Neural Networks in a Prospective Data Set of 1204 Pigmented Skin Lesions Involving Patients' Perception.	Anna Sophie Jahn, Alexander Andreas Navarini, Sara Elisa Cerminara et al.	2022
2. Evaluation de la performance d'une forme d'AI dans l'analyse de lésions cutanées				
a) Performance d'IA complexe	4	Assessing the generalisability of Deep Learning models trained on standardised and non-standardised images and their performance against teledermatologists: retrospective comparative study.	Ayooluwatomiwa I Olorunoba, Tine Vestergaard, Toan D Nguyen et al.	2022
	10	Effectiveness of an image-analysing artificial intelligence-based digital health technology to diagnose nonmelanoma skin cancer and benign skin lesions.	Helen Marsden, Caroline Morgan, Stephanie Austin et al.	2023
	13	Implementation of artificial intelligence for the detection of cutaneous melanoma within a primary care setting: prevalence and types of skin cancer in outdoor enthusiasts.	Ian J. Miller, Michael Stapelberg, Nedeljka Rosic et al.	2023
	14	Lesion identification and malignancy prediction from clinical dermatological images.	Meng Xia, Meenal K. Kheterpa, Samantha C. Wong et al.	2022
	15	MobileSkin: Classification of Skin Lesion Images Acquired Using Mobile Phone-Attached Hand-Held Dermoscopes.	Abdurrahim Yilmaz, Gulsum Gencoglan, Rahmetullah Varol et al.	2022
	21	Validation of artificial intelligence prediction models for skin cancer diagnosis using dermoscopy images: the 2019 International Skin Imaging Collaboration Grand Challenge.	Marc Combalia, Noel Codella, Veronica Rotemberg et al.	2022
b) Performance d'IA et analyse d'images séquentielles	9	Early melanoma diagnosis with sequential dermoscopic images.	Zhen Yu, Jennifer Nguyen, Toan D. Nguyen et al.	2022
	16	Monitoring patients at risk for melanoma: May convolutional neural networks replace the strategy of sequential digital dermoscopy?	Julia K. Winkler, Philipp Tschandl, Ferdinand Toberer et al.	2021

CATÉGORIE D'ARTICLE	N°	TITRE DE L'ARTICLE	AUTEURS	ANNÉE
c) Performance d'IA et aide à la décision pour la prise en charge	2	A reinforcement learning model for AI-based decision support in skin cancer.	Catarina Barata, Veronica Rotemberg, Noel C. F. Codella et al.	2023
	12	Implementation of artificial intelligence algorithms for melanoma screening in a primary care setting.	Mara Giavina-Bianchi, Raquel Machado de Sousa, Vitor Zago de Almeida Paciello et al.	2021
	24	Improving skin cancer management with ARTificial intelligence: A pre-post intervention trial of an artificial intelligence system used as a diagnostic aid for skin cancer management in a real-world specialist dermatology setting.	Claire Felmingham, Samantha MacNamara, William Cranwell et al.	2021
	19	Prospective validation of dermoscopy-based open-source artificial intelligence for melanoma diagnosis (PROVE-AI study).	Michael A. Marchetti, Emily A. Cowen, Nicholas R. Kurtansky et al.	2023
3. Performance d'une IA versus médecins (dermatologues et/ou médecins généralistes et/ou autres intervenants en soins primaires)	3	Artificial Intelligence vs Medical Providers in the Dermoscopic Diagnosis of Melanoma.	Jane M. Anderson, Izhaar Tejani, Tory Jarman et al.	2023
	11	Exploring the potential of artificial intelligence in improving skin lesion diagnosis in primary care.	Anna Escalé-Besa, Oriol Yélamos, Josep Vidal-Alaball et al.	2023
	20	Superiority of artificial intelligence compared to dermatologists and primary care providers in the diagnosis of malignant melanoma.	Jane M. Anderson, Izhaar Tejani, Tory Jarman et al.	2022
4. Evaluation de la coopération entre l'IA et les médecins (dermatologues, médecins généralistes ou autres)	5	Assessment of Diagnostic Performance of Dermatologists Cooperating With a Convolutional Neural Network in a Prospective Clinical Study: Human With Machine.	Julia K. Winkler, Andreas Blum, Katharina Kommos et al.	2023
	23	Augmenting the accuracy of trainee doctors in diagnosing skin lesions suspected of skin neoplasms in a real-world setting: A prospective controlled before-and-after study.	Young Jae Kim, Jung-Im Na, Seung Seog Han et al.	2022
	6	Clinical utility of an elastic scattering spectroscopy device in assisting primary care physician's detection of skin cancers.	Erik Jaklitsch, Todd Thames, Thomaz de Campos Silva et al.	2023
	7	Convolutional neural network assistance significantly improves dermatologists' diagnosis of cutaneous tumours using clinical images.	Wei Ba, Huan Wu, Wei W. Chen et al.	2022
	8	Diagnostic performance of augmented intelligence with 2D and 3D total body photography and convolutional neural networks in a high-risk population for melanoma under real-world conditions: A new era of skin cancer screening?	Sara E. Cerminara, Phil Cheng, Lisa Kostner et al.	2023
	25	Evaluation of Artificial Intelligence-Assisted Diagnosis of Skin Neoplasms: A Single-Center, Paralleled, Unmasked, Randomized Controlled Trial.	Seung Seog Han, Young Jae Kim, Ik Jun Moon et al.	2022

a) Analyse des lésions cutanées grâce à l'IA intégrée dans des applications pour Smartphones :

Les quatre articles suivants partageaient un objectif commun : améliorer le diagnostic des lésions cutanées, en particulier la détection des lésions malignes, en utilisant des technologies d'IA intégrées dans des applications pour smartphones.

Ou et al. (65) et Smak Gregoor et al. (66) ont, tous deux, développé des modèles utilisant des données multimodales pour améliorer la classification des lésions cutanées à partir d'images de smartphones. Ou et al. (65) ont conçu un modèle de fusion multimodale en utilisant deux encodeurs pour intégrer des images cliniques et des métadonnées (telles que l'âge du patient, l'utilisation ou non de pesticides, l'emplacement de la lésion, le type de peau...). Ils ont utilisé la base de données PAD-UPES-20 qui comprend 2 298 images représentant 6 types de lésions cutanées. Leur étude a montré que l'inclusion des métadonnées a significativement amélioré les performances du modèle, passant d'une AUROC de 0,853 à 0,947 (indiquant une excellente capacité à distinguer les lésions bénignes des lésions malignes) et d'une précision de 0,726 à $0,768 \pm 0,022$, surpassant les performances d'autres méthodes.

Smak Gregoor (66) et al. ont mené une étude de faisabilité pilote visant à évaluer l'intégration d'une application *mHealth* basée sur l'IA pour le diagnostic du cancer de la peau dans un cadre de soins primaires. Ils ont inclus 50 patients avec des lésions suspectes et l'application a été testée à la fois par les patients eux-mêmes avant leur consultation et par les médecins généralistes. Ces derniers ont ensuite ajusté leur plan de traitement en fonction des résultats de l'application. Pour vérifier son efficacité, toutes les photos prises à partir de celle-ci ont été évaluées par un minimum de trois télédermatologues qui ont analysé chaque image sans se concerter. L'histopathologie a été utilisée comme référence pour le diagnostic final. Dans les cas où elle n'était pas disponible, le diagnostic clinique du dermatologue était utilisé. L'étude a révélé que 84 % des patients ont utilisé l'application avec succès. Celle-ci a correctement identifié 80 % des lésions bénignes comme étant à faible risque et 90,9 % des lésions (pré)malignes comme étant à haut risque. Notons que 54 % des patients ayant reçu une évaluation à faible risque pour une lésion bénigne ont indiqué qu'ils auraient annulé leur visite chez le médecin généraliste, faisant confiance aux résultats de l'application. Un aspect clé de cette étude était l'intégration des retours d'expérience des utilisateurs (patients et médecins généralistes). Des données qualitatives, tels que des journaux audios, des observations de chercheurs et de groupes de discussion, ont permis de mieux comprendre les perceptions et l'acceptabilité de l'application dans un contexte de soins primaires.

Jahn et al. (67) ont testé l'efficacité diagnostique de l'application SkinVision® pour la stratification du risque de mélanome, comparant ses performances à celles des dermatologues et de dispositifs de

photographie corporelle totale (TBP) en 2D et 3D, sur un échantillon de 114 patients incluant 1 204 lésions. L'évaluation du diagnostic par l'application a été comparée aux résultats obtenus par un groupe de dermatologues et, si disponible, à l'histopathologie. Selon qu'elle était comparée à l'évaluation combinée des dermatologues et de l'IA (sensibilité plus faible) ou aux résultats histopathologiques (sensibilité plus élevée), ils ont montré que l'application SkinVision® avait une sensibilité (capacité à identifier correctement les lésions malignes, minimisant ainsi les faux négatifs) variant entre 41,3 % et 83,3 % et une spécificité (capacité du test à identifier correctement les lésions bénignes, minimisant ainsi les faux positifs) entre 60,0 % et 82,9 %. En comparaison, les dermatologues évalués par rapport à l'histopathologie, ont montré une spécificité plus élevée de 92,7 % et une sensibilité moyenne de 83 % pouvant atteindre 100 % pour les plus expérimentés. Ces résultats suggèrent que SkinVision® risquerait d'entraîner une surdétection et des excisions inutiles, limitant ainsi son utilité clinique en tant qu'outil autonome.

Kränke et al. (68) ont évalué la précision diagnostique de deux nouveaux algorithmes d'IA (« *analyze* » et « *detect* ») intégrés dans des smartphones pour détecter le potentiel malin des lésions cutanées. Le diagnostic de l'algorithme était considéré comme correct s'il correspondait à celui de deux dermatologues ou à l'histologie (si disponible). Leur étude impliquait 238 patients et 1 171 lésions. Ils ont montré que l'algorithme « *detect* » avait une précision diagnostique de 88 % (proportion de diagnostics corrects parmi tous les diagnostics réalisés), surpassant les 82 % de l'algorithme « *analyze* ». La spécificité de « *detect* » était de 94,85 % (traduisant une faible proportion de faux positifs) contre 90,32 % pour « *analyze* ». Ces algorithmes pourraient se montrer efficaces pour dépister précocement le cancer de la peau.

b) Evaluation de la performance d'une forme d'AI dans l'analyse de lésions cutanées :

Les articles suivants avaient pour principal objectif l'évaluation de la performance de formes d'Intelligence Artificielle pour diagnostiquer des lésions mélanocytaires et non mélanocytaires. Devant leur nombre important, nous avons décidé de les regrouper en plusieurs sous-catégories en regard de leurs caractéristiques. La première est constituée de six articles traitant de l'évaluation de la performance de systèmes d'IA complexes. La deuxième, composée de deux articles, porte sur l'évaluation des performances de l'IA dans l'analyse d'images séquentielles. La troisième sous-catégorie, quant à elle, regroupe quatre articles dont la particularité est de proposer une prise en charge adaptée en fonction du diagnostic établi par l'IA.

2-a) IA complexe :

Oloruntoba et al. (69) ont effectué une étude dont l'objectif est d'évaluer la performance d'un réseau de neurones convolutifs (CNN) entraîné à partir d'images standardisées en comparaison avec un CNN ayant été exercé sur des images non standardisées. Pour cela, trois modèles de CNN d'architectures identiques ont été respectivement entraînés sur trois bases de données différentes d'images standardisées ou non. Le CNN standardisé numéro 1 a été entraîné sur une base de 117 475 images de MoleMap, le CNN standardisé numéro 2 sur une base de 25 331 images de MoleMap et le CNN non standardisé sur 25 331 images de ISIC2020. Leurs performances ont ensuite été testées sur trois bases de données, une Danoise se composant de 569 images, la base de données publique ISIC2020 (33 126 images) et de données de l'Université du Queensland (422 images). Pour finir, les performances des CNN ont été comparées à celles de télédermatologues sur la base de données danoise.

Sur ces 3 tests, les performances du CNN-NS étaient inférieures à celles des deux autres groupes de CNN standardisés :

Base de données	AUROC		
	CNN-NS	CNN standardisé 1	CNN standardisé 2
Danoise	0,759	0,861	0,831
ISIC2020	0,763	0,828	0,815
Université du Queensland	0,850	0,876	0,864

Comparaison avec les télédermatologues sur la base de données Danoise :

Base de données Danoise	Dermatologues	CNN-NS	CNN standardisé 1	CNN standardisé 2
Spécificité	79,2 %	45 %	72 %	62 %
Sensibilité	82,9 %	56,3 %	74,7 %	71 %

Les modèles de CNN entraînés sur des images standardisées ont réalisé de meilleures performances que ceux entraînés sur des images non standardisées. Leur développement permettrait donc une meilleure généralisation à la classification de lésions non connues de la base de données d'entraînement.

Marsden et al. (70) ont réalisé une étude prospective dont le but est de démontrer l'efficacité d'un dispositif technologique digital d'analyse d'image basé sur l'IA dans l'identification des cancers cutanés non mélanocytaires. Les lésions ont été photographiées individuellement avec 3 smartphones différents (iPhone 6S, iPhone11, Samsung 10) équipés d'une loupe de dermoscopie DL1, puis les images ont été transférées électroniquement et analysées par le dispositif AlaMD (AI as a Medical Device). Le diagnostic de référence était basé sur l'histopathologie dans le cas des lésions biopsiées ou sur l'expertise de dermatologues lorsque celle-ci n'était pas disponible.

Au total, 611 images de lésions cutanées suspectes, dont 395 biopsiées, provenant de 572 patients, ont été photographiées, donnant les résultats suivants :

Appareil utilisé pour les photographies	AUROC pour IAaMD	
	Carcinomes épidermoïdes	Carcinomes basocellulaires
iPhone 6S	0.88	0.87
<ul style="list-style-type: none"> • Sensibilité • Spécificité 	98% 38%	94 % 28%
Samsung	0.85	0.87
iPhone 11	0.88	0.89

Les 16 lésions diagnostiquées comme mélanomes avaient été reconnues par l'IAaMD comme telles. Elle est donc capable de détecter les lésions non mélanocytaires et pré-malignes avec une précision similaire à celle de dermatologues.

L'étude de Miller et al. (71) est centrée sur la prévalence des cancers cutanés chez les sportifs en Australie, ce qui n'était pas l'objet de notre revue de la littérature. Son objectif secondaire, évaluer la performance d'un dispositif d'IA, disponible dans le commerce (*Fotofinder*), dans le diagnostic des mélanomes nous paraissait pertinent. Nous l'avons donc conservée parmi les articles retenus. Les 423 participants ont été examinés par un docteur ayant une expertise dans le dépistage des cancers cutanés. Il a sélectionné les lésions d'intérêt à soumettre au dispositif d'IA. Des 48 lésions suspectes identifiées cliniquement, le dispositif d'IA n'en a classé comme malignes que 23. Les sensibilité, spécificité et exactitude du dispositif, dans cette étude, ont été évaluées et ont obtenu respectivement une valeur de 53,33 %, 54,44 % et 54,17 %. L'AUC de cette étude a été calculée à 0,540. Concernant ce dispositif d'IA, les résultats étaient en dessous de ce qu'indiquait la littérature jusqu'à ce jour.

Xia et al. (72) ont mené une étude en deux temps dans le but d'évaluer la performance d'un système d'IA qu'ils ont développé permettant d'identifier et de classer des lésions cutanées sur des images de vue d'ensemble. Tout d'abord, l'objectif est d'identifier, à partir de photographies macroscopiques, toutes les lésions cutanées présentes sur le cliché puis de les classer en fonction de leur malignité. Pour évaluer la performance de cette technique, 488 images de patients du centre médical universitaire de Duke ont été classées par l'IA en comparant ses diagnostics avec ceux de 3 dermatologues dont les années d'expertise diffèrent. Les lésions suspectées de malignité ont été biopsiées et confirmées par l'histopathologie.

Les prédictions sur la malignité ont obtenu une AUROC de 0,790 sur l'ensemble des lésions, de 0,789 pour les lésions provenant d'images de smartphones et de 0,791 pour les images dermoscopiques. Les résultats retrouvés ont été jugés par les auteurs comme satisfaisants et de performance comparable avec ceux des dermatologues. Cependant, des limitations telles que le nombre restreint de types de

lésions ainsi que des biais dans les données pour Xia et al. (par exemple : une majorité d'images provenant de personnes à la peau claire ou moyenne) empêchent la généralisation des résultats, nécessitant une validation clinique supplémentaire pour confirmer l'efficacité de ce modèle dans des contextes réels.

L'étude de Yilmaz et al. (73) avait pour but d'évaluer la performance de systèmes d'apprentissage profond pour classer différentes lésions dermoscopiques acquises par l'utilisation de smartphones. La base de données utilisée est constituée d'images dermoscopiques exclusivement prises par des téléphones portables, soit un total de 1 688 images dermoscopiques collectées entre 2017 et 2021. Ainsi, quatre modèles d'apprentissage profond basés sur les architectures de *MobileNetV1*, *MobileNetV2*, *NASNetMobile* et *Xception* ont été développés pour identifier 8 types de lésions. Chacune d'entre elles a été classée en kératose actinique, kératose séborrhéique, lésion vasculaire, dermatofibrome, carcinome basocellulaire, carcinome épidermoïde, naevus ou mélanome. Ces images ont été recueillies par des dermatologues ayant plus de 20 ans d'expertise et tous les mélanomes validés par résultat d'histopathologie. Les modèles *MobileNetV2* et *Xception* sont les dispositifs ayant démontré la plus grande exactitude avec respectivement 89,18 et 89,64 % de réussite aux tests. Les modèles *MobileNetV1* et *NASNetMobile* montraient une performance moindre, de respectivement 76,96 % et 77,21 %.

Ces données montraient que ces dispositifs d'architecture d'apprentissage profond ont une précision élevée pour classer les lésions cutanées, se rapprochant de celle de dermatologues expérimentés. Ceux-ci pourraient être utilisés dans des situations cliniques où l'accès à des dispositifs coûteux comme les dermoscopes digitaux est difficile et fournir une aide précieuse aux cliniciens non spécialisés ainsi qu'aux dermatologues.

L'étude de Combalia et al. (74) rapportait les résultats du défi de l'ISIC 2019, dont l'objectif était de valider des modèles de prédiction basés sur l'IA pour le diagnostic de la peau à partir d'images dermoscopiques. 129 algorithmes ont été soumis par 64 équipes afin d'être testés sur un échantillon de 8 238 images. Les deux algorithmes les plus performants utilisaient des associations d'architecture d'*EfficientNet*, alors que le troisième utilisait une architecture fondée sur le ResNet. L'algorithme le plus performant a réalisé un pourcentage de 58 % sur le BCN20000, base de données la plus représentative d'un scénario clinique réaliste en comparaison avec les 82 % obtenus sur le HAM10000. Dans la plupart des catégories, à l'exception des kératoses actiniques et des lésions non incluses, les algorithmes ont été plus performants que les dermatologues. Concernant les 25 algorithmes les plus performants, 47,1 % des images ne correspondaient pas à des catégories incluses dans les bases de données

d'entraînement. Elles ont été diagnostiquées comme des lésions malignes, ce qui conduirait à une augmentation du nombre de biopsies.

Cette étude a démontré que l'exactitude des méthodes de classifications d'IA actuelles diminuait de plus de 20 % lorsqu'elle était testée sur des bases de données faites pour simuler au mieux les réalités cliniques d'une utilisation dans le monde réel.

2-b) IA et analyse d'images séquentielles :

Les deux études suivantes visaient à évaluer les performances des modèles d'IA, plus précisément des réseaux de neurones convolutifs, dans la détection précoce du mélanome. L'étude de Yu et al. (75) cherchait à déterminer si l'analyse d'images dermoscopiques séquentielles, avec un modèle de réseau spatio-temporel (intégrant les changements dynamiques des lésions), peut surpasser les performances des cliniciens expérimentés dans la détection précoce du mélanome. De son côté, Winkler et al. (76) explorait la possibilité de remplacer la stratégie de la dermoscopie numérique séquentielle (SDD) par des CNN pour détecter des mélanomes initialement identifiés sur la base de changements dynamiques.

Yu et al. (75) ont utilisé un modèle spatio-temporel pour analyser 730 images dermoscopiques provenant de 179 lésions cutanées sur 122 patients à haut risque. Le modèle, basé sur l'architecture *ResNet-34* pré-entraînée, a été testé en utilisant des images séquentielles pour capturer les changements temporels des lésions. Les résultats ont été comparés à ceux de 12 dermatologues expérimentés évaluant les mêmes images sans connaissance des diagnostics préalables. L'objectif était de comparer la sensibilité, la spécificité et la précision des deux approches (modèle et cliniciens).

Winkler et al. (76) ont mené une étude rétrospective sur 59 patients à haut risque, comprenant un total de 236 lésions (59 mélanomes et 177 nævus). Deux CNN ont été testés, CNN-1 (*GoogleNet Inception v4*) et CNN-2 (*ResNet34*), pour classifier les lésions à partir d'ensembles d'images de base (correspondant aux premières photos prises lors de l'examen initial) et d'images de suivi (séquentielles SDD). Chaque ensemble comprenait une image de mélanome détecté par des changements dynamiques lors de la SDD et trois images de nævus bénins. Les résultats des CNN ont été comparés à ceux de 26 dermatologues qui ont évalué les mêmes ensembles d'images de base.

Dans l'étude de Yu et al. (75), le modèle spatio-temporel a surpassé les dermatologues en termes de précision diagnostique, y compris dans la détection précoce des mélanomes dès la première image.

Performances	Dermatologues	Modèle spatio-temporel
Précision diagnostique	54.33 %	63.69 %
Identification précoce des mélanomes dès la première image	32,7 %	60,7 %

Dans l'étude de Winkler et al. (76), les CNN ont montré des performances inférieures à celles des dermatologues dans la classification des images de base de chaque ensemble. Cependant, après le suivi de la SDD, les sensibilités des CNN se sont améliorées. Une collaboration entre les dermatologues et le CNN-1 a permis d'améliorer la sensibilité globale (toujours pour l'analyse des images de base), mais avec une spécificité réduite (les auteurs n'en précisant pas la significativité).

Performances	Dermatologues	CNN-1	CNN-2	Collaboration CNN-1 /Dermatologues
Classification des images de base (% des ensembles correctement classés)	40,7 %	15,3 %	13,6 %	
Classification après suivi SDD (% des ensembles correctement classés)		28,8 %	20,3 %	
Sensibilité	66.1%			78,2 %
Spécificité	55,4 %			54,2 %

2-c) IA et aide à la décision pour la prise en charge :

Ces quatre études partageaient un objectif commun : évaluer l'efficacité de l'IA dans l'aide au diagnostic du cancer de la peau, en particulier pour les lésions suspectes de mélanome, son impact potentiel sur leur gestion et la prise de décision clinique.

L'étude de Barata et al. (77) avait pour objectif de déterminer si les préférences humaines, en particulier celles des dermatologues, pouvaient améliorer la précision des systèmes d'IA utilisés pour le diagnostic du cancer de la peau, en les intégrant dans un modèle d'apprentissage par renforcement (RL). Elles se traduisaient par l'attribution de poids influençant le modèle RL, permettant à celui-ci d'accorder une importance plus ou moins grande aux caractéristiques associées aux mélanomes ou aux lésions bénignes, en fonction des conséquences des erreurs de diagnostic. L'étude visait à tester la performance de ce modèle et à comparer ses résultats avec un modèle d'apprentissage supervisé (SL) classique. Dans cet objectif, l'équipe a d'abord entraîné le modèle SL sur un ensemble de 10 015 images comprenant des cancers de la peau (mélanomes et carcinomes basocellulaires), une condition précancéreuse (kératose actinique) et plusieurs conditions bénignes. Le modèle RL, basé sur le modèle SL, a été ajusté en fonction des récompenses et des pénalités définies par des experts dermatologues, tenant compte des conséquences des erreurs de diagnostic. Les deux modèles ont ensuite été testés sur un ensemble indépendant de 1 511 images et 89 dermatologues ont évalué ces images avec et sans le soutien de l'IA.

Giavina-Bianchi et al. (78) ont développé deux algorithmes d'IA pour des images cliniques (prises avec un smartphone) et dermoscopiques et les ont testés dans un cadre de soins primaires, comparant les performances diagnostiques de l'IA aux décisions des médecins généralistes. Ces algorithmes ont été

entraînés sur un ensemble de 20 388 images dermoscopiques et 19 740 images cliniques avant d'être testés dans des conditions réelles (le nombre exact de patients inclus n'est cependant pas mentionné).

Marchetti et al. (79) ont validé l'exactitude d'un algorithme open-source (ADAE) pour le diagnostic du mélanome à partir d'images dermoscopiques, tandis que Felmingham et al. (80) ont testé un CNN pour la classification des lésions cutanées en bénignes, incertaines ou malignes dans un cadre spécialisé, tout en analysant secondairement son impact sur les décisions des internes en dermatologie. Marchetti et al. (79) ont mené une étude clinique prospective sur 435 patients avec 603 lésions, dont 95 mélanomes, comparant les résultats de l'algorithme ADAE aux diagnostics histopathologiques et aux évaluations des dermatologues. Felmingham et al. (80) ont réalisé une étude avant-après intervention sur 214 patients totalisant 743 lésions ; l'histopathologie a été utilisée comme référence lorsque disponible, à défaut, l'évaluation du dermatologue traitant a servi de référence.

Les résultats de Barata et al. (77) ont montré que le modèle RL a significativement amélioré la sensibilité pour le mélanome, passant de 61,4 % avec le modèle SL à 79,5 % avec le modèle RL et de 79,6 % à 87,1 % pour le carcinome basocellulaire. L'IA a également permis une réduction de la surconfiance dans les prédictions incorrectes. Les dermatologues ont d'abord évalué les images sans assistance, puis avec l'aide des modèles d'IA, ce qui a permis de comparer directement l'impact de chaque modèle sur leurs performances. Lors de cette évaluation, le taux de diagnostics corrects et la proportion de décisions de gestion optimales (excision, thérapie locale ou surveillance) sont améliorés.

Evaluation	Sans IA	Modèle SL	Modèle RL
Taux de diagnostics corrects	68 %	75,3 %	79,9 %
Décisions de gestion optimales	57,4 %	61,7 %	65,3 %

Giavina-Bianchi et al. (78) ont obtenu une sensibilité de près de 90 % et une spécificité allant jusqu'à 89 % pour les algorithmes d'IA appliqués à des images dermoscopiques. Pour les images cliniques, la sensibilité était de 84 % et la spécificité de 82 %. Ils ont montré que l'IA a amélioré la stratification des risques dans les soins primaires, avec une sensibilité de 85 % pour l'ensemble des décisions de gestion concernant les lésions suspectes de mélanome. Celles-ci ont inclus la surveillance des lésions bénignes, la référence des patients à un spécialiste ou la décision d'effectuer une biopsie. Dans cette étude, l'IA a permis d'identifier plus précisément les lésions cutanées présentant un risque modéré à élevé, facilitant ainsi une meilleure orientation des patients vers des examens supplémentaires ou des biopsies lorsqu'une lésion suspecte de malignité était détectée. Cette assistance a contribué à une prise en charge plus efficace des patients, en évitant des interventions inutiles sur les lésions bénignes et en réduisant le risque de sous-diagnostic.

Marchetti et al. (79) ont rapporté que l’algorithme ADAE avait une sensibilité de 96,8 % et une spécificité de 37,4 %. Les dermatologues, après avoir été exposés aux scores de l'IA, ont vu leur capacité à évaluer les risques de mélanome s'améliorer, avec une augmentation de l’AUROC de 0,7798 à 0,8161, conduisant à un bénéfice clinique net et permettant de réduire les biopsies inutiles tout en maintenant une sensibilité élevée.

Dans l'étude de Felmingham et al. (80), les résultats concernant la performance de l'IA ont montré que le CNN avait une AUROC de 0,837, légèrement supérieure à celle des télédermatologues (0,807), avec un taux de faux négatifs inférieur (3,7 % contre 9,3 %). Le CNN a également amélioré la cohérence des décisions des internes, les alignant davantage avec celles des dermatologues, leur AUROC passant de 0,847 à 0,879. Cependant, il a aussi augmenté les biopsies de lésions bénignes mais non significativement ($p = 0,12$) du point de vue statistique.

c) Performance d'une IA versus médecins (dermatologues et/ou médecins généralistes et/ou autres intervenants en soins primaires)

Les deux études regroupées dans ce paragraphe avaient pour objectif commun la comparaison des performances de médecins dermatologues, généralistes et autres acteurs de santé à celles d’une IA.

Anderson et al. (81,82) ont réalisé une étude publiée dans deux journaux différents sous deux titres distincts :

- Superiority of Artificial Intelligence in the Diagnostic Performance of Malignant Melanoma Compared to Dermatologists and Primary Care Providers en 2022 ;
- Artificial Intelligence vs Medical Providers in the Dermoscopic Diagnosis of Melanoma en 2023.

Cette étude avait pour objectif d’évaluer la performance de l’application d’IA *Triage* par rapport aux acteurs de soins primaires et secondaires afin de déterminer son utilité pour dépister les cancers cutanés et réduire le temps avant diagnostic. Cette application est le résultat de l’association d’un CNN pré-entraîné et d’un système d’apprentissage par renforcement. 100 images dermoscopiques, dont 80 bénignes et 20 mélanomes vérifiés par biopsie, ont été analysées par l’application mais également par 23 dermatologues, 7 médecins généralistes, et 12 paramédicaux (IDE, IPA et assistants médicaux).

Résultats obtenus par l’étude :

Evaluation	IA	Dermatologues	Médecins généralistes	Paramédicaux
Sensibilité	80 %	77 %	78 %	61 %
Spécificité	95 %	57 %	41 %	53 %
Exactitude	92 %	61 %	48 %	55 %
Valeur prédictive positive	80 %	32 %	25 %	26 %
Valeur prédictive négative	98 %	92 %	90 %	85 %

L'application *Triage* avait donc obtenu de manière significative une plus grande exactitude, sensibilité et spécificité que l'ensemble des intervenants médicaux, dont l'exactitude varie de 32 à 78 % selon la catégorie d'intervenant. Cependant, bien que la spécificité, l'exactitude et les valeurs prédictives positive et négative aient été plus élevées chez l'IA, il n'y avait pas de différence significative entre l'application et les dermatologues en termes de sensibilité ou lors de l'identification des lésions malignes.

Escalé-Besa et al. (83) ont effectué une étude dont l'objectif était de valider un modèle d'apprentissage automatique comme outil d'aide au diagnostic pour des lésions cutanées de diverses pathologies dermatologiques (44 pathologies différentes) dans un environnement de soins primaires. Pour cela, les performances du modèle ont été comparées à celles des 11 médecins généralistes et des 3 télédermatologues qui acceptèrent d'y participer dans un contexte clinique. Les 100 patients inclus ont été recrutés dans 6 centres de santé primaire dans la région de Catalogne en Espagne et consultaient leur médecin généraliste pour évaluer une lésion cutanée. Celles-ci ont été photographiées et anonymisées avant d'être intégrées dans le modèle d'interface Autoderm afin d'obtenir un diagnostic par l'IA. Les résultats ont été analysés selon la précision du modèle de Machine Learning (ML) dénommée Top-1, Top-3 et Top-5. Top-1 signifie que le diagnostic correct correspondait à la première proposition du modèle, le Top-3 signifie qu'il faisait partie des trois premières propositions et le Top-5 qu'il était parmi les cinq premiers.

Résultats obtenus par l'étude :

Exactitude Diagnostique	Modèle machine learning	Télédermatologues	Médecins généralistes
Top-1 Toutes lésions	0,39	0,72	0,64
Top-3 Toutes lésions	0.61	0.9	0.76
Top-3 Lésions bénignes	0,87	0,84	0,76
Top-3 tumeurs malignes	0,67		0,92

L'exactitude de diagnostic du modèle de d'apprentissage automatique augmentait significativement lorsque l'on incluait le Top-3.

Du fait de sa configuration, cette étude apporte des avancées sur l'application de l'IA dans la pratique routinière d'un médecin généraliste.

d) Évaluation de la coopération entre l'IA et les médecins (dermatologues, médecins généralistes ou autres)

Les six études suivantes partageaient un objectif commun : évaluer la coopération entre l'IA et les médecins pour améliorer le diagnostic des lésions cutanées suspectes. Plus précisément, Winkler et al. (84), Kim et al. (85) et Jaklitsch et al. (86) visaient à évaluer si l'IA pouvait améliorer la précision

diagnostique des médecins non spécialisés et des dermatologues, particulièrement dans des contextes cliniques réels pour les néoplasmes cutanés. Ba et al. (87), Cerminara et al. (88) et Han et al. (89) visaient également à tester l'efficacité de la coopération IA-médecins, mais en ciblant particulièrement l'amélioration des performances diagnostiques dans des populations à haut risque ou des environnements de soins spécialisés.

Winkler et al. (84) ont mené une étude clinique prospective impliquant 22 dermatologues et 228 lésions chez 188 patients. Les dermatologues ont d'abord réalisé des diagnostics à l'œil nu, puis avec dermoscopie. Ces images dermoscopiques ont été évaluées par l'IA basée sur un CNN (*Moleanalyzer Pro*). La performance diagnostique des dermatologues sans assistance a été comparée à celle qu'ils obtiennent en utilisant le CNN.

Kim et al. (85) ont réalisé une étude prospective contrôlée impliquant 285 patients présentant des lésions cutanées suspectes de malignité. Les patients ont été examinés par des médecins en formation, dont des internes en médecine générale et des résidents en dermatologie. Les médecins ont d'abord examiné les lésions et proposé un diagnostic, avec ou sans l'assistance d'un algorithme d'IA (type IA non spécifié). L'IA assistait les médecins en fournissant des suggestions de diagnostics probables. Les performances diagnostiques des médecins, avec et sans assistance IA, ont été comparées aux diagnostics histopathologiques lorsque disponibles. Sur les 285 patients, 270 cas ont été confirmés par biopsie et analyse histopathologique. Pour les 15 autres cas qui présentaient des lésions clairement bénignes, sans nécessité de biopsie, le diagnostic final a été établi par des dermatologues expérimentés après un examen clinique approfondi.

Jaklitsch et al. (86) ont évalué l'utilité d'un dispositif portable basé sur la spectroscopie de diffusion élastique (ESS pour *Elastic Scattering Spectroscopy*, une technique d'analyse optique non invasive permettant d'examiner la structure des tissus en étudiant la lumière diffusée), combinée à un modèle d'IA pour aider les médecins généralistes à diagnostiquer et à prendre en charge les lésions cutanées. Ils ont conduit une étude observationnelle prospective où 57 médecins généralistes ont évalué 50 lésions de 44 patients. Les médecins les ont d'abord évaluées de manière traditionnelle (uniquement à partir de l'examen visuel), puis avec l'aide de l'ESS et les résultats des deux approches ont ensuite été comparés pour évaluer l'amélioration apportée par l'IA.

Ba et al. (2022) (87) ont utilisé un CNN (*EfficientNet-B3*) pour assister 18 dermatologues dans le diagnostic de 400 images issues d'un ensemble de 2 107 lésions cutanées chez 143 patients. Les performances des dermatologues, avec et sans assistance IA, ont été comparées.

Cerminara et al. (88) ont évalué un modèle IA basé sur la photographie corporelle totale (TBP) en 2D (*FotoFinder ATBM Master*) et en 3D (*Vectra WB360*) pour 1 690 lésions chez 143 patients à haut risque

de mélanome. Ces derniers ont bénéficié d'une cartographie corporelle complète à l'aide de ces systèmes, assistés par un CNN, pour évaluer les scores de risque de malignité. Les résultats des évaluations par IA ont ensuite été comparés aux diagnostics cliniques réalisés par des dermatologues et aux analyses histopathologiques lorsqu'une biopsie était réalisée.

Han et al. (89) ont mené un essai clinique randomisé portant sur 576 cas de lésions cutanées suspectes, impliquant des internes en médecine générale et des internes en dermatologie. L'étude a comparé la précision diagnostique des médecins lorsqu'ils étaient assistés par une IA (*Model Dermatology Build2020*, basée sur un CNN) à celle obtenue lorsqu'ils réalisaient les diagnostics sans assistance. Dans le groupe assisté par l'IA, les internes devaient d'abord formuler leurs diagnostics seuls, puis, ils recevaient des suggestions de l'IA (incluant les diagnostics potentiels et un score de malignité) pour affiner leur évaluation finale. Le test de référence pour le diagnostic des lésions provenait soit d'une analyse histopathologique (524 biopsies), soit, pour les lésions clairement bénignes (52 cas), d'un diagnostic établi par des dermatologues expérimentés.

Winkler et al. (84) ont trouvé que l'assistance par le CNN a amélioré la sensibilité des dermatologues de 84,2 % à 100 %, et la spécificité de 72,1 % à 83,7 %, ce qui a permis une réduction des excisions inutiles et une meilleure précision diagnostique. Cette amélioration a été particulièrement marquée chez les dermatologues moins expérimentés, mais ceux qui l'étaient davantage ont également bénéficié de l'assistance de l'IA. De même, Ba et al. (87) ont observé que la précision diagnostique des dermatologues avec l'assistance de l'IA augmentait de 62,78 % à 76,60 %, la sensibilité de 83,21 % à 89,56 % et la spécificité de 80,92 % à 87,90 %.

Kim et al. (85) ont rapporté une amélioration significative de la précision diagnostique des médecins en formation avec l'IA : la précision Top-1 (le diagnostic principal proposé est correct dès le premier essai) est passée de 58,3 % avec l'IA contre 46,5 % sans l'IA. Cette dernière a permis une meilleure détection des lésions malignes et a contribué à réduire les excisions inutiles. Han et al. (89) ont également montré que l'IA a amélioré la précision des médecins non spécialisés (internes en médecine générale) de 30,7 % à 54,7 %, bien que cela ne soit pas significatif pour les internes en dermatologie qui n'obtiennent qu'une hausse légère de la précision qui passe de 51,7 % à 53,1 %. Le modèle IA a montré une performance autonome équivalente à celle des dermatologues pour la détection des malignités (sensibilité de 66,7 %, spécificité de 90,6 %).

Cerminara et al. (88) ont constaté que le CNN 3D a obtenu une sensibilité de 90 %, des résultats comparables à ceux des dermatologues (sensibilité de 90 %), mais avec une spécificité plus faible (64,6 % contre 92,3 % pour les dermatologues), tandis que le CNN 2D était moins performant. Le modèle 3D a également montré une meilleure reproductibilité pour les évaluations de risque, fournissant des

résultats plus cohérents et fiables lors des mesures répétées, avec un coefficient de corrélation de 0,89 contre 0,79 pour le modèle 2D.

Enfin, Jaklitsch et al. (86) ont montré que l'IA testée a augmenté la sensibilité diagnostique des médecins généralistes de 67 % à 88 %, mais a légèrement réduit la spécificité de 53 % à 40 %. Cela indique que l'IA a permis une détection plus efficace des cancers de la peau, mais au prix d'une augmentation de faux positifs, entraînant potentiellement plus d'excisions inutiles. L'IA a amélioré la gestion des lésions suspectes en soins primaires, en augmentant la sensibilité des décisions cliniques (biopsie ou référence à un dermatologue) de 81 % à 94 %, tout en renforçant la confiance des cliniciens en leurs décisions.

IV - DISCUSSION

Après analyse exhaustive de la littérature ciblant les articles publiés entre août 2021 et août 2023, nous atteignons notre objectif principal en confirmant que l'IA pourrait effectivement aider les médecins généralistes à diagnostiquer précocement les lésions malignes dermatologiques dans le cadre d'une coopération homme-machine. Les études menées en environnements cliniques contrôlés ont, pour la plupart, montré que l'IA améliore la précision diagnostique en augmentant à la fois la sensibilité et la spécificité dans la détection des cancers de la peau. Les résultats de certains articles portent à croire qu'elle aurait le potentiel d'influencer positivement la gestion des patients en optimisant les décisions de biopsie ou de référence, notamment en réduisant les excisions inutiles, ce qui concourt à maintenir une sensibilité élevée et à limiter les risques de sous-diagnostic.

Cependant, bien que les résultats théoriques soient prometteurs, aucune étude n'a encore évalué l'impact de l'IA à grande échelle ou de manière systématique dans la pratique clinique. Des limitations subsistent, telles que la nécessité de généraliser les résultats à des populations plus diversifiées ou de poursuivre la validation de ces outils dans des environnements de soins primaires. Ces résultats seront discutés plus en détail ci-dessous.

Les quatre études traitant du potentiel de l'IA pour améliorer le diagnostic des lésions cutanées via des applications pour smartphones ont montré des avancées notables sans occulter pour autant des limitations importantes.

Ou et al. (65) ont démontré que l'intégration de métadonnées améliore significativement la précision diagnostique de leur modèle de fusion multimodale. Cependant, leur étude n'a pas encore été validée en milieu clinique direct et le nombre limité de types de lésions dans leur base de données restreint la généralisation des résultats. Smak Gregoor et al. (66) ont montré que leur application *mHealth* pourrait être utilisée efficacement dans les soins primaires pour identifier les lésions nécessitant une attention médicale plus poussée. Toutefois, l'utilisation de l'application doit encore être améliorée car une proportion significative de patients n'a pas pu l'utiliser sans assistance. Kränke et al. (68) ont montré que l'algorithme « *detect* » pourrait être efficace pour le dépistage des lésions cutanées suspectes, surclassant l'autre algorithme testé « *analyse* ». Jahn et al. (67) ont trouvé, quant à eux, que l'application SkinVision® pourrait entraîner une surdétection de lésions malignes préoccupante si elle est utilisée en tant qu'outil unique. Leur étude est aussi limitée par le petit nombre de mélanomes confirmés par histopathologie, affaiblissant la robustesse des conclusions. En outre, les études de Jahn et al. (67) et de Kränke et al. (68) ont été réalisées auprès de populations à haut risque de mélanome. La spécificité de ces cohortes pourrait biaiser les résultats et limiter leur généralisation à des

populations moins à risque. La variabilité de la qualité des images capturées par les smartphones dans l'étude de Kränke et al. (68) constitue également une limitation importante.

Les études dont l'objectif était d'éprouver des techniques ou algorithmes d'IA ont montré, dans la majorité des résultats, une performance comparable à celle des dermatologues, à l'exception de l'étude de Miller et al. Cet article traitait de la prévalence des mélanomes chez les sportifs en Australie. Les auteurs justifiaient cette différence de résultats par rapport au reste de la littérature par les limitations intrinsèques de l'étude. Parmi elles, on comptait un nombre restreint de lésions étudiées ainsi qu'une population spécifique qui se différenciait de celle dans laquelle le dispositif *Fotofinder* avait originellement été testé de par une exposition cutanée aux UV importante.

En effet, les résultats en faveur d'une performance similaire ou supérieure aux dermatologues sont à considérer avec prudence en regard des diverses limitations que comportent les études les rapportant. La plupart d'entre elles n'ont pas été réalisées dans un environnement clinique strict, notamment celles de Oloruntoba et al. (69), Combalia et al. (74) et Yilmaz et al. (73). Cela entraîne un biais en faveur des systèmes d'IA et nous laisse douter de la performance de ces dispositifs sortis d'un contexte théorique. Les études de Marsden et al. (70), Yilmaz et al. (73) et de Xia et al. (72) comptent également comme limitation une faible variabilité des phénotypes inclus dans les images testées, ce qui limite de manière conséquente la généralisation des résultats à une population différente. A cela s'ajoute, pour l'étude de Xia et al. (72), un nombre de lésions étudiées réduit, limitant une fois de plus la validité externe de l'étude. Les études de Yilmaz et al. (73) et de Miller et al. (71) ont exclu les images de lésions se trouvant dans des endroits rendant les photographies difficiles (cou, plis, cuir chevelu...) mais également celles des muqueuses et des ongles. Nous tenions à mentionner que l'article de Oloruntoba et al. (69), portant sur la supériorité d'un entraînement par images standardisées, signalait qu'elles étaient de bien meilleure qualité que celles non standardisées, biaisant l'étude de manière significative. La performance du CNN pouvait en effet être améliorée par une base de données de meilleure qualité, sans qu'elle soit standardisée pour autant.

Il est important de remarquer que, dans l'article de Combalia et al. (74), les algorithmes conçus pour identifier des images étrangères à leur base d'entraînement performaient moins bien que les experts, ce qui, à ce jour, met en évidence de sérieux soucis sur la sécurité de leur utilisation dans le cadre d'une pratique clinique. De plus, un certain nombre de lésions non identifiées étaient classées en malignes, ce qui augmente le nombre de biopsies injustifiées réalisées.

Les deux études de Yu et al. (75) et Winkler et al. (76) s'inscrivent dans le cadre de l'amélioration des outils de diagnostic précoce du mélanome, en testant les capacités de modèles d'IA, notamment des réseaux de neurones convolutifs, pour détecter des mélanomes à leur stade initial. Les recherches

montrent que, dans des contextes spécifiques (analyse d'images séquentielles ou surveillance à long terme), les CNN peuvent offrir une aide précieuse aux cliniciens en améliorant la détection des mélanomes. Toutefois, les études soulignent qu'à ce stade, ils ne sont pas encore en mesure de remplacer totalement les méthodes cliniques traditionnelles telle que la dermoscopie numérique séquentielle. Une approche collaborative entre les dermatologues et les modèles d'IA semble être la voie la plus prometteuse pour améliorer la sensibilité diagnostique tout en maintenant un taux raisonnable de spécificité. En pratique, l'intégration de l'IA pourrait permettre une surveillance plus efficace des patients à haut risque, réduire les excisions inutiles tout en assurant une détection précoce des mélanomes.

Néanmoins, ces recherches présentent des limitations communes, comme le fait qu'elles ont été menées sur des échantillons relativement restreints de patients (122 dans l'étude de Yu et 59 dans celle de Winkler), ce qui limite la généralisation des résultats. En outre, aucune des deux études n'a intégré les CNN de manière prospective dans des décisions cliniques réelles, empêchant ainsi une évaluation complète de leur impact potentiel sur la gestion clinique en temps réel. Plus spécifiquement, dans l'étude de Yu et al. (75), le modèle a montré une performance moindre pour les petites lésions ou celles ayant des bords mal définis et le contexte clinique simplifié des médecins évaluateurs (pas d'accès aux informations complètes sur les antécédents des patients ou autres éléments contextuels habituellement disponibles) pourrait affecter la comparaison avec la pratique réelle. L'étude de Winkler et al., quant à elle, n'a pas inclus l'évaluation des images de suivi par les dermatologues, rendant la comparaison avec les CNN moins exhaustive. De plus, leurs CNN testés n'étaient pas conçus pour intégrer les changements dynamiques des lésions cutanées, un aspect clé dans la surveillance des mélanomes.

Les études de Giavina-Bianchi et al. (78), Marchetti et al. (79), Felmingham et al. (80) et Barata et al. (77) montrent que l'IA, sous différentes formes (algorithmes supervisés, CNN, modèles d'apprentissage par renforcement), apporte une contribution significative à l'amélioration des diagnostics de cancers de la peau et des décisions de prise en charge du patient. Giavina-Bianchi et al. (78) ont constaté que l'IA a amélioré la précision des décisions de gestion en soins primaires, facilitant l'accès aux soins spécialisés. Marchetti et al. (79) ont observé que l'algorithme ADAE permettait de réduire les biopsies inutiles tout en maintenant une sensibilité élevée, améliorant ainsi l'efficacité clinique. De leur côté, Felmingham et al. (80) ont noté que l'utilisation d'un CNN a aidé à aligner les décisions des résidents avec celles des dermatologues expérimentés, bien que cela ait conduit à une légère augmentation des biopsies de lésions bénignes. Barata et al. (77) ont démontré que l'apprentissage par renforcement (RL), en intégrant les préférences humaines, a amélioré non seulement la sensibilité des diagnostics,

mais aussi la prise en charge clinique en aidant les dermatologues à prendre des décisions plus optimales.

Une limitation partagée par toutes les études traitant de la performance de l'IA et de l'aide à la prise de décision réside dans la nécessité de validations supplémentaires en pratique clinique. Giavina-Bianchi et al. (78) et Felmingham et al. (80) soulignent l'importance d'étendre les tests à des populations plus diversifiées, notamment en ce qui concerne les types de peau ou la localisation des lésions. Felmingham et al. (80) relèvent également un potentiel biais d'échantillonnage lié aux cliniques spécialisées très fréquentées (où les patients présentent généralement un risque plus élevé de développer un cancer de la peau). Ces deux points pourraient limiter la généralisation des résultats à d'autres contextes cliniques, notamment en soins primaires. Barata et al. (77) mentionnent que l'intégration des préférences des patients, en plus de celles des dermatologues, pourrait encore améliorer la pertinence clinique de l'IA. En outre, leur modèle d'apprentissage par renforcement nécessite un réentraînement constant, ce qui peut compliquer sa mise en œuvre dans des contextes réels. Enfin, les études de Felmingham et al. (80) et Marchetti et al. (79) mettent en avant la question des biopsies inutiles, soulignant le besoin d'un meilleur équilibre entre sensibilité accrue et gestion des cas bénins. Marchetti et al. (79) se heurtent également à la limitation d'un échantillon monocentrique et restreint en nombre de dermatologues, ainsi qu'à l'absence d'évaluation sur des phototypes plus variés.

Les deux études comparant la performance de médecins dermatologues, généralistes et autres paramédicaux à celle de l'IA ont montré des résultats en partie discordants. Anderson et al. (81,82) concluent à une plus grande spécificité, sensibilité et exactitude de l'application d'IA Triage par rapport aux cliniciens, dermatologues compris. Escalé-Besa et al. (83), en revanche, retrouvent une sensibilité et une exactitude de l'IA inférieure à celle des dermatologues et des médecins généralistes.

Ces résultats contradictoires peuvent être expliqués par les limitations intrinsèques à ces deux articles, dont les principales sont partagées. La première de celles-ci se fonde sur la taille réduite d'images analysées, en limitant la généralisation des résultats à une population plus large. La deuxième limite que ces études ont en commun porte sur le fait que la population étudiée soit issue d'une zone géographique restreinte et bien définie. Les phénotypes de peau les moins courants en dépistage de carcinomes cutanés et mélanomes n'y sont donc pas représentés. L'étude de Escalé-Besa et al. (83) contient également un biais de performance du côté des médecins généralistes évaluateurs, les volontaires étant majoritairement des médecins intéressés par la dermatologie. Il est probable qu'ils aient donc une formation, dans ce domaine, plus élaborée que la moyenne de la population dont ils sont issus et que l'on souhaite évaluer. Anderson et al. (81,82) se posent également la question de

l'importance qu'ont, sur les résultats de leur étude, les compétences des différents professionnels de santé en dermoscopie. Poursuivre de nouvelles études dans un milieu clinique réel, avec une restriction de ces biais et sur un nombre plus important de patients, paraît d'une importance cruciale avant de conclure à une performance suffisante de l'IA et d'en valider l'utilisation dans la pratique clinique quotidienne des intervenants de santé.

Les recherches menées sur la coopération entre les médecins et l'IA présentent des avancées importantes et des points communs notables. Toutes démontrent que l'IA améliore la précision diagnostique, particulièrement chez les médecins moins expérimentés, permettant de mieux détecter les lésions malignes. Une autre avancée majeure est que l'IA, en augmentant la sensibilité diagnostique pour les cancers de la peau, contribue à réduire les excisions inutiles et à améliorer la gestion des lésions suspectes. Ces bénéfices sont observés tant dans les modèles utilisant des CNN que dans d'autres approches comme l'ESS.

Cependant, des différences et des limitations existent dans ces études. Cerminara et al.(88) ont montré que le CNN 2D était nettement moins performant que le CNN 3D et les dermatologues. C'est une observation spécifique à cette étude, d'autres n'ayant pas relevé de différences aussi marquées entre les performances des différentes approches d'IA. De leur côté, Han et al.(89) ont constaté que l'IA n'améliorait pas significativement les performances des internes en dermatologie, contrairement aux études de Winkler et al. (84) et Ba et al. (87) qui ont montré une amélioration des performances diagnostiques même chez les dermatologues expérimentés.

En ce qui concerne les limitations, plusieurs études partagent des aspects similaires. Par exemple, la population d'étude était souvent réduite à des phototypes cutanés spécifiques. Han et al. (89) et Kim et al. (85) ont mené leurs recherches principalement sur des patients asiatiques (essentiellement de phototypes III et IV), tandis que Winkler et al. (84) ont travaillé avec des patients de phototypes II et III, ce qui restreint la généralisation des résultats à d'autres populations. Winkler et al. (84) ont également mis en avant une focalisation sur les lésions mélanocytaires et un environnement contrôlé, ce qui réduit la portée de leurs conclusions dans des conditions cliniques réelles. De surcroît, certains articles, comme celui de Jaklitsch et al. (86), ont noté une diminution de la spécificité avec l'IA, augmentant le risque de faux positifs et donc d'excisions inutiles. Enfin, Han et al. (89) et Kim et al. (85) ont également mentionné le biais potentiel lié à l'entraînement de l'IA sur des images spécifiques, ce qui pourrait limiter son efficacité face à des cas atypiques en dehors de ce cadre.

Notre premier objectif secondaire était de recenser les différentes méthodes d'IA existantes dans l'analyse des lésions cutanées. Les 25 articles analysés montrent une diversité de méthodes utilisées

pour le diagnostic des lésions cutanées. L'une des approches les plus couramment employées est celle des réseaux de neurones convolutifs (CNN). Par exemple, Yilmaz et al. ont utilisé des architectures CNN telles que *MobileNet* et *EfficientNet* pour classifier des images dermoscopiques, permettant une analyse précise des lésions en identifiant des motifs complexes dans les images. Giavina-Bianchi et al. ont également utilisé *EfficientNet* pour le diagnostic des lésions cutanées, tout en ayant recours à un réseau adversaire génératif (GAN) pour concevoir des images artificielles de mélanome afin de compenser le déséquilibre des données et améliorer l'entraînement de leur modèle. Barata et al. ont comparé un modèle d'apprentissage supervisé (SL) à un modèle d'apprentissage par renforcement (RL), en ajustant les décisions du modèle RL en fonction des retours des dermatologues pour améliorer la sensibilité et la spécificité du diagnostic des lésions cutanées. Han et al. et Felmingham et al. ont, quant à eux, utilisé des modèles d'apprentissage profond (Deep Learning) basés sur des CNN pour analyser des images dermoscopiques et cliniques, contribuant ainsi à la détection des cancers de la peau et à une meilleure gestion des patients.

Le second était de déterminer le coût de l'IA, et son accessibilité par rapport à l'intervention de spécialistes. De manière générale, bien que l'IA soit souvent perçue comme une technologie prometteuse capable de réduire les coûts et d'améliorer l'accès aux soins, il n'existe pas de chiffres concrets quant aux investissements nécessaires à son déploiement en soins primaires. Plusieurs études soulignent les avantages potentiels de l'IA, notamment pour alléger la charge de travail des spécialistes et optimiser la gestion des cas complexes, tout en permettant un diagnostic plus rapide et précis dans des contextes où les ressources sont limitées. L'accessibilité accrue des smartphones, combinée à l'utilisation de technologies d'IA, est vue comme une solution efficace pour rendre le dépistage des cancers de la peau plus abordable. C'est particulièrement vrai dans les zones rurales ou les pays en développement car les dermatoscopes coûtent cher et ne sont pas toujours disponibles. Cependant, des obstacles persistent, tels que le coût élevé de mise en œuvre des technologies dans les soins primaires, la nécessité d'une formation spécialisée pour les praticiens ainsi que la confiance limitée des cliniciens et des patients dans la fiabilité des applications d'IA. Malgré ces défis, l'amélioration du diagnostic précoce par l'IA pourrait non seulement augmenter les taux de survie des patients, mais aussi réduire les coûts de santé à long terme, en évitant les traitements plus coûteux des cancers avancés.

Bien que les résultats de notre analyse systématique de la littérature apportent des éléments intéressants sur l'efficacité et l'accessibilité de l'IA en dermatologie, il est important de souligner plusieurs biais méthodologiques qui pourraient influencer l'interprétation de ces données.

L'un des principaux biais de cette étude est que de nombreuses technologies d'IA et d'apprentissage automatique (AA) développées à des fins commerciales ne font pas l'objet de publications dans des revues universitaires, ce qui signifie qu'elles n'ont pas été identifiées dans notre recherche systématique issue de 3 bases de données. Contrairement à l'étude de *The Lancet* sur laquelle nous nous appuyons, nous n'avons pas effectué de recherche manuelle dans les archives ou réseaux commerciaux (Google, Microsoft, Apple, IBM, etc.), ce qui aurait pu enrichir et compléter notre analyse. De la même manière, n'y sont pas non plus incluses des données de la littérature grise (thèse ou mémoire, rapports d'agences gouvernementales, documents de travail...).

De plus, nos recherches ont révélé peu d'algorithmes d'IA ou d'AA spécifiquement développés ou testés dans des contextes de soins primaires. Bien que nous nous soyons concentrées sur des études menées en milieu clinique, il y avait souvent une prépondérance d'études impliquant des dermatologues plutôt que des médecins généralistes ou d'autres praticiens de soins primaires. Cela peut biaiser les résultats en faveur des comparateurs de l'IA, les dermatologues étant généralement plus expérimentés dans la détection des lésions cutanées. Le peu d'études évaluant les médecins généralistes les associaient souvent à d'autres acteurs de soins primaires (non formés au diagnostic de lésions cutanées), affaiblissant significativement leurs performances en comparaison avec ceux de l'IA.

Un autre biais réside dans le fait que deux des articles inclus dans ce travail exploitent les résultats d'une même étude. Ils ont été rédigés par les mêmes auteurs et publiés à un an d'intervalle dans deux revues différentes. Afin d'éviter d'attribuer un poids excessif à des résultats redondants, nous n'avons pris en compte ces données qu'une seule fois dans notre analyse.

Enfin, bien que certaines études aient été menées dans le contexte des soins primaires, il n'y en avait pas suffisamment dans ce cadre spécifique pour réaliser des analyses statistiques complètes et pertinentes sur l'impact de l'IA dans ces environnements. Cela limite notre capacité à généraliser nos conclusions sur l'efficacité de l'IA dans les soins primaires.

Notre étude présente plusieurs forces méthodologiques importantes. Elle s'appuie, tout d'abord, sur une recherche exhaustive de la littérature en utilisant trois grandes bases de données : MEDLINE, Embase et Web of Science. Les équations de recherche utilisées sont fondées sur celles validées dans l'étude d'O.T. Jones, publiée dans *The Lancet*, et intègrent des termes MeSH validés par des experts. Cela garantit une rigueur et une exhaustivité des résultats, assurant que toutes les sources pertinentes sont prises en compte.

Ensuite, les articles ont été sélectionnés de manière rigoureuse, seuls ceux publiés dans des revues académiques à comité de lecture ont été retenus. Cela permet de garantir que chaque étude incluse a fait l'objet d'une évaluation par des pairs, d'en assurer la qualité scientifique et l'impartialité des données exploitées.

Nous avons aussi choisi d'exclure les revues systématiques de la littérature et les méta-analyses afin d'éviter une sur-représentation de certains résultats. Cette décision permet d'assurer une évaluation plus équilibrée et fidèle des données, sans distorsion induite par des études déjà synthétisées.

Un autre point fort réside dans la double évaluation de chaque article sélectionné, tant sur la qualité du rapport que sur la méthodologie utilisée. Des outils d'évaluation standardisés et reconnus, tels que les grilles STROBE, TREND, CONSORT, QUADAS-2, MINORS et Jadad, ont été employés, garantissant ainsi une analyse rigoureuse, transparente et fiable.

L'originalité de notre étude constitue également une force significative. L'importance et la pertinence de notre travail dans le domaine est renforcé puisqu'après une recherche approfondie dans la littérature grise, il apparaît qu'aucune autre thèse n'aborde spécifiquement ce sujet. En effet, lors de notre recherche sur SUDOC (le catalogue du Système Universitaire de Documentation), nous avons trouvé une thèse portant sur le sujet de l'IA en médecine générale. Son objectif principal était de comprendre si les généralistes rencontraient des difficultés à diagnostiquer les tumeurs cutanées non pigmentées et à orienter leurs patients vers des dermatologues ainsi que d'évaluer si un outil de diagnostic de ces tumeurs utilisant l'IA pouvait être perçu comme utile dans la pratique quotidienne (90).

A cette occasion, nous avons consulté les sites de l'OMS, de la Société Française de Dermatologie (SFD) et de la Haute Autorité de Santé (HAS). Bien que nous n'ayons pas identifié d'études spécifiques sur l'évaluation de l'intelligence artificielle (IA) appliquée au diagnostic des lésions cutanées, plusieurs rapports de ces institutions fournissent des recommandations importantes sur l'usage de l'IA en santé. Les rapports de l'OMS (publié le 28 juin 2021) (91,92), de la HAS (validé le 9 novembre 2022) (93) et de la SFD (daté du 16 septembre 2022) (94) soulignent des points convergents.

Ils insistent sur la nécessité d'encadrer l'utilisation de l'IA afin de garantir qu'elle reste un outil complémentaire, utilisé sous la supervision des professionnels de santé. Ils mettent en garde contre les risques liés à l'automatisation excessive, notamment l'absence d'accompagnement humain lors du diagnostic et l'impact sur la relation patient-médecin. La responsabilité médicale reste un enjeu central, l'IA ne devant pas remplacer la décision clinique.

Les trois agences partagent également des préoccupations autour des biais algorithmiques, de la nécessité d'une validation continue des outils d'IA et de l'importance d'une transparence sur leur fonctionnement. Le rapport de la HAS met en avant la difficulté de valider ces technologies en vie réelle et la rareté des études prospectives robustes. De son côté, la SFD recommande une formation accrue des professionnels à l'usage de l'IA, tout en appelant à une réglementation stricte pour prévenir les dérives potentielles, comme l'utilisation de l'IA par le grand public sans supervision médicale.

Ainsi, ces institutions s'accordent à dire que l'IA, en dermatologie et en soins premiers, peut offrir des opportunités pour améliorer le diagnostic des lésions cutanées, mais elles insistent sur un encadrement rigoureux et une utilisation sous contrôle médical pour maximiser les bénéfices tout en minimisant les risques.

Enfin, la gestion des conflits d'intérêts a été soigneusement prise en compte. Bien que notre étude n'ait pas inclus les réseaux commerciaux dans la recherche de la littérature (ce qui peut entraîner une certaine perte de données), ce choix permet de limiter l'influence des conflits d'intérêts potentiels. Il est cependant à noter que certains articles inclus mentionnent des financements privés, nous l'avons pris en compte de manière transparente.

La conception de cette étude a suivi les recommandations PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), visant à garantir la transparence et l'exhaustivité dans la présentation d'une revue systématique ou d'une méta-analyse. En appliquant à notre étude la grille d'évaluation PRISMA, nous avons obtenu un score de 17/27 (Annexe n° 5). Nous avons également utilisé la grille de notation AMSTAR (A Measurement Tool to Assess systematic Reviews), pour apprécier la validité interne de l'étude en évaluant la qualité méthodologique et la note retrouvée était de 7/11 (Annexe n° 5).

Par ailleurs, les deux auteurs déclarent n'avoir aucun conflit d'intérêts et aucun financement spécifique n'a été reçu pour la rédaction de cette revue systématique de la littérature.

V - CONCLUSION

Notre revue de la littérature permet d'actualiser les avancées effectuées par les techniques d'IA dans le diagnostic et le dépistage des cancers cutanés mélanocytaires et non mélanocytaires. Ces technologies se développent à grande vitesse, cependant, elles restent, à ce jour, insuffisamment éprouvées dans un contexte clinique. Des études supplémentaires de grande envergure sont nécessaires afin de vérifier leur sécurité d'utilisation dans le monde réel. Par ailleurs, l'utilisation de l'IA dans la pratique clinique soulève également des questions éthiques et d'engagement de la responsabilité juridique. C'est notamment le cas concernant la prise de décision médicale automatisée et la répartition des responsabilités entre l'IA et les professionnels de santé en cas d'erreurs diagnostiques.

le 07/10/2024
Vu, la Présidente du jury
Pr Julie DUPOUY
Médecine Générale

Toulouse, le 08/10/2024

Vu et permis d'imprimer
La Présidente de l'Université Toulouse III – Paul Sabatier
Faculté de Santé
Par délégation,
Le Doyen-Directeur
Du Département de Médecine, Maïeutique, et Paramédical
Professeur Thomas GEERAERTS



VI - RÉFÉRENCES :

1. Institut National du Cancer. Panorama des cancers. 2023 p. 1-21.
2. Ben Kacem A. La pratique de la dermatologie en médecine générale : recours au spécialiste et besoins de formation en Picardie [thèse]. Amiens Univ Picardie. 2019;
3. Arnault F. Approche territoriale des spécialités médicales et chirurgicales. 2023 janv p. 1-316.
4. Gassia V, Grognard C, Michaud T. La consultation esthétique au sein de l'activité du dermatologue. Ann Dermatol Vénérologie. oct 2009;136:S390-2.
5. OMG - Observatoire de la Médecine Générale. Informations épidémiologiques sur les pathologies et leur prise en charge en ville [Internet]. 2009. Disponible sur: <http://omg.sfm.org/>
6. Letrilliart L, Supper I, Schuurs M, Darmon D, Boulet P, Favre M, et al. ECOGEN : étude des Eléments de la COnsultation en médecine GENérale. 1 janv 2014;
7. Collège des enseignants en dermatologie de France. Référentiel des Collèges de Dermatologie. 8ème Edition. Elsevier Masson; 2022.
8. Bernard G, Alain D, Marc P, Géraldine J, Elif H, Thomas J, et al. ACTUALISATION DES DONNEES CONCERNANT LE MELANOME STADE III : NOUVELLES RECOMMANDATIONS DU GROUPE DE CANCEROLOGIE CUTANEE. 2018;1-27.
9. Jones OT, Matin RN, Schaar M van der, Bhayankaram KP, Ranmuthu CKI, Islam MS, et al. Artificial intelligence and machine learning algorithms for early detection of skin cancer in community and primary care settings: a systematic review. Lancet Digit Health. 1 juin 2022;4(6):e466-76.
10. Ring C, Cox N, Lee JB. Dermatoscopy. Clin Dermatol. août 2021;39(4):635-42.
11. Kittler H, Pehamberger H, Wolff K, Binder M. Diagnostic accuracy of dermoscopy. Lancet Oncol. mars 2002;3(3):159-65.
12. Thomas L, Braun RP. Atlas de dermoscopie. Elsevier Masson; 2007. 232 p.
13. Dermoscopy Atlas | Diagnosis Detail [Internet]. Disponible sur: <https://www.dermoscopyatlas.com/case/314/2706>
14. Giacco-Demoulins N. La dermatoscopie en médecine générale [Internet]. 2020. Disponible sur: <https://www.larevuedupraticien.fr/article/la-dermatoscopie-en-medecine-generale>
15. Braun RP, Rabinovitz HS, Oliviero M, Kopf AW, Saurat JH. Dermoscopy of pigmented skin lesions. J Am Acad Dermatol. janv 2005;52(1):109-21.
16. Marghoob A, Braun R, Kopf AW. An Atlas of Dermoscopy. CRC Press; 2004. 408 p.
17. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal. déc 2017;42:60-88.
18. Laino ME, Cancian P, Politi LS, Della Porta MG, Saba L, Savevski V. Generative Adversarial Networks in Brain Imaging: A Narrative Review. J Imaging. 23 mars 2022;8(4):83.

19. Møllersen K, Kirchesch H, Zortea M, Schopf TR, Hindberg K, Godtliebsen F. Computer-Aided Decision Support for Melanoma Detection Applied on Melanocytic and Nonmelanocytic Skin Lesions: A Comparison of Two Systems Based on Automatic Analysis of Dermoscopic Images. *BioMed Res Int*. 2015;2015:579282.
20. Perednia DA. What dermatologists should know about digital imaging. *J Am Acad Dermatol*. 1 juill 1991;25(1):89-108.
21. Russell, S., Norvig, P. *Artificial Intelligence: A Modern Approach*. 4th ed. NJ USA: Prentice Hall: Hoboken; 2020.
22. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys*. 1 déc 1943;5(4):115-33.
23. International Business Machines Corporation. Qu'est-ce qu'un réseau de neurones ? | IBM [Internet]. Disponible sur: <https://www.ibm.com/fr-fr/topics/neural-networks>
24. Zemouri R, Devalland C, Valmary-Degano S, Zerhouni N. Intelligence artificielle : quel avenir en anatomie pathologique ? *Ann Pathol*. 1 avr 2019;39(2):119-29.
25. Erickson BJ, Korfiatis P, Kline TL, Akkus Z, Philbrick K, Weston AD. Deep Learning in Radiology: Does One Size Fit All? *J Am Coll Radiol JACR*. mars 2018;15(3 Pt B):521-6.
26. McCorduck Pamela. *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence* [Internet]. 2nd éd. USA: AK Peters Ltd Taylor & Francis Group; 2004. 598 p. Disponible sur: <https://www.routledge.com/Machines-Who-Think-A-Personal-Inquiry-into-the-History-and-Prospects-of-Artificial-Intelligence/McCorduck/p/book/9781568812052>
27. Widrow B, Lehr MA. 30 years of adaptive neural networks: perceptron, Madaline, and backpropagation. *Proc IEEE*. sept 1990;78(9):1415-42.
28. Lodwick GS, Keats TE, Dorst JP. THE CODING OF ROENTGEN IMAGES FOR COMPUTER ANALYSIS AS APPLIED TO LUNG CANCER. *Radiology*. août 1963;81:185-200.
29. Mitchell, T. *Machine Learning* [Internet]. 1st ed. New York, NY, USA: McGraw-Hill Education Ltd.; 1997. Disponible sur: <http://www.cs.cmu.edu/~tom/mlbook.html>
30. Technologies D. What is a Support Vector Machine? [Internet]. *Datatron*. 2021 [cité 12 sept 2024]. Disponible sur: <https://datatron.com/what-is-a-support-vector-machine/>
31. Melbin K, Raj YJV. Integration of modified ABCD features and support vector machine for skin lesion types classification. *Multimed Tools Appl*. 1 mars 2021;80(6):8909-29.
32. Alamri A, Alsaed D. On the development of a skin cancer computer aided diagnosis system using support vector machine. *Biosci Biotechnol Res Commun*. 25 juin 2019;12(2):297-308.
33. Afifi S, GholamHosseini H, Sinha R. SVM classifier on chip for melanoma detection. In: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) [Internet]. Seogwipo: IEEE; 2017 [cité 12 sept 2024]. p. 270-4. Disponible sur: <https://ieeexplore.ieee.org/document/8036814/>
34. Karalis G. *Decision Trees and Applications*. *Adv Exp Med Biol*. 2020;1194:239-42.

35. Tanaka T, Voigt MD. Decision tree analysis to stratify risk of de novo non-melanoma skin cancer following liver transplantation. *J Cancer Res Clin Oncol*. 1 mars 2018;144(3):607-15.
36. Sugahara S, Ueno M. Exact Learning Augmented Naive Bayes Classifier. *Entropy Basel Switz*. 20 déc 2021;23(12):1703.
37. Alwan O.F. Skin cancer images classification using naïve bayes. *Emergent J Educ Discov Lifelong Learn*. 2022;3:19-29.
38. Balaji P, Hung BT, Chakrabarti P, Chakrabarti T, Elngar AA, Aluvalu R. A novel artificial intelligence-based predictive analytics technique to detect skin cancer. *PeerJ Comput Sci*. 24 mai 2023;9:e1387.
39. Mobiny A, Singh A, Van Nguyen H. Risk-Aware Machine Learning Classifier for Skin Lesion Diagnosis. *J Clin Med*. 17 août 2019;8(8):1241.
40. Zhang Z. A gentle introduction to artificial neural networks. *Ann Transl Med*. oct 2016;4(19):370.
41. Divya D, Ganeshbabu T. Fitness adaptive deer hunting-based region growing and recurrent neural network for melanoma skin cancer detection. *Int J Imaging Syst Technol*. 15 mars 2020;30.
42. Sivaraj S, Malmathanraj R, Palanisamy P. Detecting anomalous growth of skin lesion using threshold-based segmentation algorithm and Fuzzy K-Nearest Neighbor classifier. *J Cancer Res Ther*. 2020;16(1):40-52.
43. Oukil S, Kasmi R, Mokrani K, García-Zapirain B. Automatic segmentation and melanoma detection based on color and texture features in dermoscopic images. *Skin Res Technol*. 15 nov 2021;28(2):203-11.
44. Gershman SJ. The Generative Adversarial Brain. *Front Artif Intell [Internet]*. 2019 [cité 7 avr 2024];2. Disponible sur: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7861215/>
45. Qin Z, Liu Z, Zhu P, Xue Y. A GAN-based image synthesis method for skin lesion classification. *Comput Methods Programs Biomed*. oct 2020;195:105568.
46. Rashid H, Tanveer MA, Aqeel Khan H. Skin Lesion Classification Using GAN based Data Augmentation. *Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Int Conf*. juill 2019;2019:916-9.
47. Ghosh A, Sufian A, Sultana F, Chakrabarti A, De D. Fundamental Concepts of Convolutional Neural Network. In 2020. p. 519-67.
48. Lambert R. Focus : Le Réseau de Neurones Convolutifs [Internet]. *Pensée Artificielle*. 2019. Disponible sur: <https://penseeartificielle.fr/focus-reseau-neurones-convolutifs/>
49. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J Big Data*. 2021;8(1):53.
50. Combalia M, Codella N, Rotemberg V, Carrera C, Dusza S, Gutman D, et al. Validation of artificial intelligence prediction models for skin cancer diagnosis using dermoscopy images: the

- 2019 International Skin Imaging Collaboration Grand Challenge. *Lancet Digit Health*. 1 mai 2022;4(5):e330-9.
51. Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data*. 14 août 2018;5(1):180161.
 52. Hernández-Pérez C, Combalia M, Podlipnik S, Codella NCF, Rotemberg V, Halpern AC, et al. BCN20000: Dermoscopic Lesions in the Wild. *Sci Data*. 17 juin 2024;11:641.
 53. Mendonca T, Ferreira PM, Marques JS, Marcal ARS, Rozeira J. PH² - a dermoscopic image database for research and benchmarking. *Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Int Conf*. 2013;2013:5437-40.
 54. Dupouy J., Kinouani S. Revue systématique de la littérature : chic et pratique ! *E-Respect*. mars 2013;19-26.
 55. Cuschieri S. The CONSORT statement. *Saudi J Anaesth*. avr 2019;13(Suppl 1):S27-30.
 56. Des Jarlais DC, Lyles C, Crepaz N, TREND Group. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health*. mars 2004;94(3):361-6.
 57. Vandembroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS Med*. 16 oct 2007;4(10):e297.
 58. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials*. févr 1996;17(1):1-12.
 59. Slim K, Nini E, Forestier D, Kwiatkowski F, Panis Y, Chipponi J. Methodological index for non-randomized studies (minors): development and validation of a new instrument. *ANZ J Surg*. sept 2003;73(9):712-6.
 60. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol*. sept 2010;25(9):603-5.
 61. Pallot A, Rostagno S. Newcastle-Ottawa Scale : traduction française de l'échelle de qualité méthodologique pour les études de cohorte. *Kinésithérapie Rev*. juill 2021;21(235):3-4.
 62. Walter FM, Thompson MJ, Wellwood I, Abel GA, Hamilton W, Johnson M, et al. Evaluating diagnostic strategies for early detection of cancer: the CanTest framework. *BMC Cancer*. 14 juin 2019;19(1):586.
 63. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *Open Med Peer-Rev Indep Open-Access J*. 2009;3(3):e123-130.
 64. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*. 15 févr 2007;7:10.

65. Ou C, Zhou S, Yang R, Jiang W, He H, Gan W, et al. A deep learning based multimodal fusion model for skin lesion diagnosis using smartphone collected clinical images and metadata. *Front Surg.* 2022;9:1029991.
66. Smak Gregoor AM, Sangers TE, Eekhof JA, Howe S, Revelman J, Litjens RJ, et al. Artificial intelligence in mobile health for skin cancer diagnostics at home (AIM HIGH): a pilot feasibility study. *EClinicalMedicine.* juin 2023;60:102019.
67. Jahn AS, Navarini AA, Cerminara SE, Kostner L, Huber SM, Kunz M, et al. Over-Detection of Melanoma-Suspect Lesions by a CE-Certified Smartphone App: Performance in Comparison to Dermatologists, 2D and 3D Convolutional Neural Networks in a Prospective Data Set of 1204 Pigmented Skin Lesions Involving Patients' Perception. *Cancers [Internet].* 7 août 2022;14(15). Disponible sur: <https://www.embase.com/search/results?subaction=viewrecord&id=L2018496755&from=export>
68. Kränke T, Tripolt-Droschl K, Röd L, Hofmann-Wellenhof R, Koppitz M, Tripolt M. New AI-algorithms on smartphones to detect skin cancer in a clinical setting-A validation study. *PloS One.* 2023;18(2):e0280670.
69. Oloruntoba I, Vestergaard T, Nguyen TD, Ge Z, Mar V. Assessing generalisability of deep learning models trained on standardised and non-standardised images. *Australas J Dermatol.* 2022;63(SUPPL 1):100.
70. Marsden H, Palamaras I, Kemos P, Greenhalgh J. Effectiveness of an image-analysing artificial intelligence-based digital health technology to diagnose nonmelanoma skin cancer and benign skin lesions. *Br J Dermatol.* 2023;188((Marsden H.; Kemos P.; Greenhalgh J.) Skin Analytics, London, United Kingdom):iv43.
71. Miller IJ, Stapelberg M, Rosic N, Hudson J, Coxon P, Furness J, et al. Implementation of artificial intelligence for the detection of cutaneous melanoma within a primary care setting: prevalence and types of skin cancer in outdoor enthusiasts. *PeerJ.* 2023;11:e15737.
72. Xia M, Kheterpal MK, Wong SC, Park C, Ratliff W, Carin L, et al. Lesion identification and malignancy prediction from clinical dermatological images. *Sci Rep.* 23 sept 2022;12(1):15836.
73. Yilmaz A, Gencoglan G, Varol R, Demircali AA, Keshavarz M, Uvet H. MobileSkin: Classification of Skin Lesion Images Acquired Using Mobile Phone-Attached Hand-Held Dermoscopes. *J Clin Med.* 30 août 2022;11(17).
74. Combalia M, Codella N, Rotemberg V, Carrera C, Dusza S, Gutman D, et al. Validation of artificial intelligence prediction models for skin cancer diagnosis using dermoscopy images: the 2019 International Skin Imaging Collaboration Grand Challenge. *Lancet Digit Health.* mai 2022;4(5):e330-9.
75. Yu Z, Ge Z, Nguyen J, Nguyen TD, Kelly J, McLean C, et al. Early melanoma diagnosis with sequential dermoscopic images. *IEEE Trans Med Imaging.* mars 2022;41(3):633-46.
76. Winkler JK, Tschandl P, Toberer F, Sies K, Fink C, Enk A, et al. Monitoring patients at risk for melanoma: May convolutional neural networks replace the strategy of sequential digital dermoscopy? *Eur J Cancer Oxf Engl 1990.* janv 2022;160:180-8.
77. Barata C, Rotemberg V, Codella NCF, Tschandl P, Rinner C, Akay BN, et al. A reinforcement learning model for AI-based decision support in skin cancer. *Nat Med.* août 2023;29(8):1941-6.

78. Giavina-Bianchi M, de Sousa RM, Paciello VZ de A, Vitor WG, Okita AL, Prôa R, et al. Implementation of artificial intelligence algorithms for melanoma screening in a primary care setting. *PloS One*. sept 2021;16(9):e0257006.
79. Marchetti MA, Cowen EA, Kurtansky NR, Weber J, Dauscher M, DeFazio J, et al. Prospective validation of dermoscopy-based open-source artificial intelligence for melanoma diagnosis (PROVE-AI study). *NPJ Digit Med*. 12 juill 2023;6(1):127.
80. Felmingham C, MacNamara S, Cranwell W, Williams N, Wada M, Adler NR, et al. Improving Skin cancer Management with ARTificial Intelligence (SMARTI): protocol for a preintervention/postintervention trial of an artificial intelligence system used as a diagnostic aid for skin cancer management in a specialist dermatology setting. *BMJ Open*. 4 janv 2022;12(1):e050203.
81. Anderson JM, Tejani I, Jarman T, Kellett L, Moy RL. Artificial Intelligence vs Medical Providers in the Dermoscopic Diagnosis of Melanoma. *Cutis*. mai 2023;111(5):254-8.
82. Anderson JM, Tejani I, Jarman T, Kellett L, Moy R. Superiority of artificial intelligence compared to dermatologists and primary care providers in the diagnosis of malignant melanoma. *J Invest Dermatol*. 2022;142(8):S108.
83. Escalé-Besa A, Yélamos O, Vidal-Alaball J, Fuster-Casanovas A, Miró Catalina Q, Börve A, et al. Exploring the potential of artificial intelligence in improving skin lesion diagnosis in primary care. *Sci Rep*. 15 mars 2023;13(1):4293.
84. Winkler JK, Blum A, Kommos K, Enk A, Toberer F, Rosenberger A, et al. Assessment of Diagnostic Performance of Dermatologists Cooperating With a Convolutional Neural Network in a Prospective Clinical Study: Human With Machine. *JAMA Dermatol*. 1 juin 2023;159(6):621-7.
85. Kim YJ, Na JI, Han SS, Won CH, Lee MW, Shin JW, et al. Augmenting the accuracy of trainee doctors in diagnosing skin lesions suspected of skin neoplasms in a real-world setting: A prospective controlled before-and-after study. *PloS One*. 2022;17(1 January):e0260895.
86. Jaklitsch E, Thames T, de Campos Silva T, Coll P, Oliviero M, Korb Ferris L. Clinical utility of an AI-powered, Handheld Elastic Scattering Spectroscopy Device on the Diagnosis and Management of Skin Cancer by Primary Care Physicians. *J Prim Care Community Health*. 2023;14(4):1-7.
87. Ba W, Wu H, Chen WW, Wang SH, Zhang ZY, Wei XJ, et al. Convolutional neural network assistance significantly improves dermatologists' diagnosis of cutaneous tumours using clinical images. *Eur J Cancer Oxf Engl* 1990. juill 2022;169:156-65.
88. Cerminara SE, Cheng P, Kostner L, Huber S, Kunz M, Maul JT, et al. Diagnostic performance of augmented intelligence with 2D and 3D total body photography and convolutional neural networks in a high-risk population for melanoma under real-world conditions: A new era of skin cancer screening? *Eur J Cancer Oxf Engl* 1990. sept 2023;190:112954.
89. Han SS, Kim YJ, Moon IJ, Jung JM, Lee MY, Lee WJ, et al. Evaluation of Artificial Intelligence-Assisted Diagnosis of Skin Neoplasms: A Single-Center, Paralleled, Unmasked, Randomized Controlled Trial. *J Invest Dermatol*. sept 2022;142(9):2353-2362.e2.
90. Samaran R. Étude d'Intérêt en Médecine Générale d'une Application d'Aide au Diagnostic des Tumeurs Cutanées Non Pigmentées via un Outil Utilisant l'Intelligence Artificielle. [SPÉCIALITÉ DERMATOLOGIE ET VÉNÉROLOGIE]. Université de Nantes; 2020.

91. World Health Organization. Ethics and governance of artificial intelligence for health [Internet]. World Health Organization, Avenue Appia 20 1121 Geneva 27 Switzerland; 2021 juin p. 150. (WHO guidance). Disponible sur: <https://www.who.int/publications/i/item/9789240029200>
92. Tarik Jasarevic, Chargé de communication, OMS, Tarik Jasarevic. L'OMS publie le premier rapport mondial sur l'intelligence artificielle (IA) appliquée à la santé et six principes directeurs relatifs à sa conception et à son utilisation. Communiqué de presse [Internet]. 28 juin 2021 [cité 11 sept 2024]; Disponible sur: <https://www.who.int/fr/news/item/28-06-2021-who-issues-first-global-report-on-ai-in-health-and-six-guiding-principles-for-its-design-and-use>
93. HAS, ROCHEREAU Aude, RENNER Simon, COLLIGNON Corinne, RENNESSON Marina, LEFEVRE Maud. Intégration des dispositifs médicaux numériques à usage professionnel dans la pratique. Haute Autorité de santé – Service communication et information 5, avenue du Stade de France – 93218 SAINT-DENIS LA PLAINE; 2022 nov p. 51. (Évaluer les technologies de santé).
94. Société Française de Dermatologie, Groupe de réflexion éthique (GÉD), Groupe de Télédermatologie et e-santé (Teldes). La Société Française de Dermatologie se mobilise pour encadrer le champ d'application de l'intelligence artificielle en dermatologie. Société Française de Dermatologie et de pathologie sexuellement transmissible (SFD) MAISON DE LA DERMATOLOGIE 10, Cité Malesherbes – 75009 Paris; 2022 sept p. 3.

VII - ANNEXES

VII- 1 ANNEXE 1 : Équations de recherche sur les trois bases de données utilisées

STRATÉGIE de RECHERCHE EMBASE (réalisée par Mme DEVRIENDT Anne-Morgane et M SIDRE Colin de l'Université Paris Cité)	
N° de la recherche	Demande
1	((early or earlier or facilitat*or aid* or augment* or improv*) adj3 (detect* or diagnos* or recogni* or identif* or classif* or decision*)).ab,ti.
2	triage.ab,ti.
3	#1 OR #2
4	melanoma.ab,ti.
5	exp *melanoma/
6	((suspicious or pigmented) adj3 (skin tumour or skin lesion)).ab,ti.
7	"basal cell carcinoma".ab,ti.
8	exp *basal cell carcinoma/
9	squamous cell carcinoma.ab,ti.
10	exp *squamous cell carcinoma/
11	skin cancer.mp. or exp *skin cancer/
12	skin neoplasm.mp. or exp *skin tumor/
13	skin malignan*.ab,ti.
14	#4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13
15	artificial intelligence.ab,ti.
16	exp *artificial intelligence/
17	neural network*.ab,ti.
18	exp *artificial neural network/
19	.ab,ti.
20	exp *machine learning/
21	((machine or computer) adj2 vision).ab,ti.
22	((unsupervised or supervised or weak* supervised or limited supervision) adj3 (learning or network or algorithm)).ab,ti.
23	exp *computer assisted diagnosis/
24	deep learning.ab,ti.
25	support vector machines.ab,ti.
26	exp *support vector machine/
27	15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26
28	3 and 14 and 27
29	limit 28 to yr="2021 -2023"

STRATÉGIE de RECHERCHE WEB of SCIENCE			
N° de la recherche	Demande	Filtres	Résultats
1	TI=((early or earlier or facilitat or aid or augment or improv) NEAR/3 (detect or diagnos or recogni or identif or classif or decision))		8 211
2	TS=((early or earlier or facilitat or aid or augment or improv) NEAR/3 (detect or diagnos or recogni or identif or classif or decision))		67 541
3	TI=trriage		9 188
4	TS=trriage		30 602
5	#4 OR #3 OR #2 OR #1		97 881
6	TI=melanoma		108 939
7	TS=melanoma		201 227
8	TI=((suspicious or pigmented) NEAR/3 (tumour or tumor or lesion))		2 172
9	TS=((suspicious or pigmented) NEAR/3 (tumour or tumor or lesion))		12 009
10	TI=(basal cell carcinoma)		9 858
11	TS=(basal cell carcinoma)		30 459
12	TI=(squamous cell carcinoma)		73 551
13	TS=(squamous cell carcinoma)		177 851
14	TI=(skin NEAR/3 (tumour or tumor or neoplasm or cancer or malignancy))		16 801
15	TS=(skin NEAR/3 (tumour or tumor or neoplasm or cancer or malignancy))		58 119
16	#15 OR #14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #8 OR #7 OR #6		427 470
17	TI=artificial intelligence OR TS=artificial intelligence		170 166
18	TI=neural network OR TS=neural network		730 436
19	TI=machine learning OR TS=machine learning		447 653
20	TI=((machine OR computer) NEAR/3 (vision)) OR TS=((machine OR computer) NEAR/3 (vision))		94 046
21	TI=((unsupervised or supervised or supervision) NEAR/3 (learning or network or algorithm)) OR TS=((unsupervised or supervised or supervision) NEAR/3 (learning or network or algorithm))		94 437
22	TI=(computer assisted diagnosis) OR TS=(computer assisted diagnosis)		8 400
23	TI=(deep learning) OR TS=(deep learning)		339 547
24	TI=(support vector machines) OR TS=(support vector machines)		139 440
25	#24 OR #23 OR #22 OR #21 OR #20 OR #19 OR #18 OR #17		1 443 213
26	#5 AND #16 AND #25	from 2021 - 2023	801

STRATÉGIE de RECHERCHE MEDLINE (Pubmed)

N° de la recherche	Demande	Filtres	Détails de la recherche	Résultats
1	((early or earlier or facilitat* or aid or augment* or improv*) AND (detect* or diagnos* or recogni* or identif* or classif* or decision*))		("early"[All Fields] OR "earlier"[All Fields] OR "facilitat*" [All Fields] OR "aid"[All Fields] OR "augment*" [All Fields] OR "improv*" [All Fields]) AND ("detect*" [All Fields] OR "diagnos*" [All Fields] OR "recogni*" [All Fields] OR "identif*" [All Fields] OR "classif*" [All Fields] OR "decision*" [All Fields])	2 822 682
2	Triage		"triage"[MeSH Terms] OR "triage"[All Fields] OR "trriages"[All Fields] OR "triaged"[All Fields] OR "triaging"[All Fields]	32 986
3	#1 OR #2		((("early"[All Fields] OR "earlier"[All Fields] OR "facilitat*" [All Fields] OR "aid"[All Fields] OR "augment*" [All Fields] OR "improv*" [All Fields]) AND ("detect*" [All Fields] OR "diagnos*" [All Fields] OR "recogni*" [All Fields] OR "identif*" [All Fields] OR "classif*" [All Fields] OR "decision*" [All Fields])) OR ("triage"[MeSH Terms] OR "triage"[All Fields] OR "trriages"[All Fields] OR "triaged"[All Fields] OR "triaging"[All Fields]))	2 845 626
4	("melanoma"[MeSH Terms]) OR melanoma		"melanoma"[MeSH Terms] OR "melanoma"[MeSH Terms] OR "melanoma"[All Fields] OR "melanomas"[All Fields] OR "melanomas"[All Fields]	159 282
5	((suspicious or pigmented) AND (skin tumour or skin lesion))		("suspicious"[All Fields] OR "suspiciously"[All Fields] OR ("pigment"[All Fields] OR "pigmental"[All Fields] OR "pigmentation"[MeSH Terms] OR "pigmentation"[All Fields] OR "pigmentations"[All Fields] OR "pigmented"[All Fields] OR "pigmenting"[All Fields] OR "pigmentous"[All Fields] OR "pigments"[All Fields])) AND ("skin neoplasms"[MeSH Terms] OR ("skin"[All Fields] AND "neoplasms"[All Fields]) OR "neoplasms"[All Fields]) OR "skin neoplasms"[All Fields] OR ("skin"[All Fields] AND "tumor"[All Fields]) OR "skin tumor"[All Fields] OR ("skin"[All Fields] AND "tumour"[All Fields]) OR "skin tumour"[All Fields] OR (("skin"[MeSH Terms] OR "skin"[All Fields]) AND ("lesion"[All Fields] OR "lesion s"[All Fields] OR "lesional"[All Fields] OR "lesions"[All Fields])))	19 633
6	((suspicious or pigmented) AND (skin tumour or skin lesion))		"carcinoma, basal cell"[MeSH Terms] OR ("carcinoma"[All Fields] AND "basal"[All Fields] AND "cell"[All Fields]) OR "basal cell carcinoma"[All Fields] OR ("basal"[All Fields] AND "cell"[All Fields] AND "carcinoma"[All Fields]) OR "carcinoma, basal cell"[MeSH Terms]	31 797
7	(Squamous cell carcinoma) OR (carcinoma, squamous cell[MeSH Terms])		"carcinoma, squamous cell"[MeSH Terms] OR ("carcinoma"[All Fields] AND "squamous"[All Fields] AND "cell"[All Fields]) OR "squamous cell carcinoma"[All Fields] OR ("squamous"[All Fields] AND "cell"[All Fields] AND "carcinoma"[All Fields]) OR "carcinoma, squamous cell"[MeSH Terms]	194 605
8	(Skin cancer) OR (skin neoplasm*[MeSH Terms])		"skin neoplasms"[MeSH Terms] OR ("skin"[All Fields] AND "neoplasms"[All Fields]) OR "skin neoplasms"[All Fields] OR ("skin"[All Fields] AND "cancer"[All Fields]) OR "skin cancer"[All Fields] OR "skin neoplasm*" [MeSH Terms]	224 855
9	skin malignan*		("skin"[MeSH Terms] OR "skin"[All Fields]) AND "malignan*" [All Fields]	53 836

STRATÉGIE de RECHERCHE MEDLINE (Pubmed)

N° de la recherche	Demande	Filtres	Détails de la recherche	Résultats
10	#4 OR #5 OR #6 OR #7 OR #8 OR #9		<p>"melanoma"[MeSH Terms] OR ("melanoma"[MeSH Terms] OR "melanoma"[All Fields] OR "melanomas"[All Fields] OR "melanoma s"[All Fields]) OR (("suspicious"[All Fields] OR "suspiciously"[All Fields] OR ("pigment"[All Fields] OR "pigmental"[All Fields] OR "pigmentation"[MeSH Terms] OR "pigmentation"[All Fields] OR "pigmentations"[All Fields] OR "pigmented"[All Fields] OR "pigmenting"[All Fields] OR "pigmentous"[All Fields] OR "pigments"[All Fields])) AND ("skin neoplasms"[MeSH Terms] OR ("skin"[All Fields] AND "neoplasms"[All Fields]) OR "skin neoplasms"[All Fields] OR ("skin"[All Fields] AND "tumor"[All Fields]) OR "skin tumor"[All Fields] OR ("skin"[All Fields] AND "tumour"[All Fields]) OR "skin tumour"[All Fields] OR ("skin"[MeSH Terms] OR "skin"[All Fields]) AND ("lesion"[All Fields] OR "lesion s"[All Fields] OR "lesional"[All Fields] OR "lesions"[All Fields])))) OR ("carcinoma, basal cell"[MeSH Terms] OR ("carcinoma"[All Fields] AND "basal"[All Fields] AND "cell"[All Fields]) OR "basal cell carcinoma"[All Fields] OR ("basal"[All Fields] AND "cell"[All Fields] AND "carcinoma"[All Fields]) OR "carcinoma, basal cell"[MeSH Terms]) OR ("carcinoma, squamous cell"[MeSH Terms] OR ("carcinoma"[All Fields] AND "squamous"[All Fields] AND "cell"[All Fields]) OR "squamous cell carcinoma"[All Fields] OR ("squamous"[All Fields] AND "cell"[All Fields] AND "carcinoma"[All Fields]) OR "carcinoma, squamous cell"[MeSH Terms]) OR ("skin neoplasms"[MeSH Terms] OR ("skin"[All Fields] AND "neoplasms"[All Fields]) OR "skin neoplasms"[All Fields] OR ("skin"[All Fields] AND "cancer"[All Fields]) OR "skin cancer"[All Fields] OR "skin neoplasm*"[MeSH Terms]) OR (("skin"[MeSH Terms] OR "skin"[All Fields]) AND "malignan*"[All Fields])</p>	512 630
11	(artificial intelligence) OR (ai artificial intelligence[MeSH Terms])		<p>"artificial intelligence"[MeSH Terms] OR ("artificial"[All Fields] AND "intelligence"[All Fields]) OR "artificial intelligence"[All Fields] OR "artificial intelligence"[MeSH Terms]</p>	201 722
12	(neural network*) OR (neural networks computer[MeSH Terms])		<p>((("neural"[All Fields] OR "neuralization"[All Fields] OR "neuralize"[All Fields] OR "neuralized"[All Fields] OR "neuralizes"[All Fields] OR "neuralizing"[All Fields] OR "neurally"[All Fields]) AND "network*"[All Fields]) OR "neural networks, computer"[MeSH Terms])</p>	155 333
13	(machine learning) OR (machine learning[MeSH Terms])		<p>"machine learning"[MeSH Terms] OR ("machine"[All Fields] AND "learning"[All Fields]) OR "machine learning"[All Fields] OR "machine learning"[MeSH Terms]</p>	119 254
14	((machine or computer) AND vision)		<p>("machine"[All Fields] OR "machine s"[All Fields] OR "machines"[All Fields] OR ("computability"[All Fields] OR "computable"[All Fields] OR "computating"[All Fields] OR "computation"[All Fields] OR "computational"[All Fields] OR "computations"[All Fields] OR "compute"[All Fields] OR "computed"[All Fields] OR "computer s"[All Fields] OR "computers"[MeSH Terms] OR "computers"[All Fields] OR "computer"[All Fields] OR "computes"[All Fields] OR "computing"[All Fields] OR "computational"[All Fields])) AND ("vision s"[All Fields] OR "vision, ocular"[MeSH Terms] OR ("vision"[All Fields] AND "ocular"[All Fields]) OR "ocular vision"[All Fields] OR "vision"[All Fields] OR "visions"[All Fields] OR "visioning"[All Fields])</p>	31 192

STRATÉGIE de RECHERCHE MEDLINE (Pubmed)

N° de la recherche	Demande	Filtres	Détails de la recherche	Résultats
15	((unsupervised or supervised or (weak* supervised) or limited supervision) AND (learning or network or algorithm))		("unsupervised"[All Fields] OR ("organization and administration"[MeSH Terms] OR ("organization"[All Fields] AND "administration"[All Fields]) OR "organization and administration"[All Fields] OR "supervision"[All Fields] OR "supervise"[All Fields] OR "supervised"[All Fields] OR "supervises"[All Fields] OR "supervising"[All Fields] OR "supervisions"[All Fields]) OR ("weak*" [All Fields] AND ("organization and administration"[MeSH Terms] OR ("organization"[All Fields] AND "administration"[All Fields]) OR "organization and administration"[All Fields] OR "supervision"[All Fields] OR "supervise"[All Fields] OR "supervised"[All Fields] OR "supervises"[All Fields] OR "supervising"[All Fields] OR "supervisions"[All Fields])) OR ("limit"[All Fields] OR "limitation"[All Fields] OR "limitations"[All Fields] OR "limited"[All Fields] OR "limiting"[All Fields] OR "limits"[All Fields]) AND ("organization and administration"[MeSH Terms] OR ("organization"[All Fields] AND "administration"[All Fields]) OR "organization and administration"[All Fields] OR "supervision"[All Fields] OR "supervise"[All Fields] OR "supervised"[All Fields] OR "supervises"[All Fields] OR "supervising"[All Fields] OR "supervisions"[All Fields])) AND ("learning"[MeSH Terms] OR "learning"[All Fields] OR "learn"[All Fields] OR "learned"[All Fields] OR "learning s"[All Fields] OR "learnings"[All Fields] OR "learns"[All Fields] OR ("network"[All Fields] OR "network s"[All Fields] OR "networked"[All Fields] OR "networker"[All Fields] OR "networkers"[All Fields] OR "networking"[All Fields] OR "networks"[All Fields]) OR ("algorithm s"[All Fields] OR "algorithmic"[All Fields] OR "algorithmically"[All Fields] OR "algorithmics"[All Fields] OR "algorithmization"[All Fields] OR "algorithms"[MeSH Terms] OR "algorithms"[All Fields] OR "algorithm"[All Fields]))	193 639
16	computer assisted diagnosis[MeSH Terms]		"diagnosis, computer assisted"[MeSH Terms]	86 716
17	(Deep Learning) OR ("deep learning"[MeSH Terms])		"deep learning"[MeSH Terms] OR ("deep"[All Fields] AND "learning"[All Fields]) OR "deep learning"[All Fields] OR "deep learning"[MeSH Terms]	59 218
18	(Support Vector Machine) OR ("support vector machine"[MeSH Terms])		"support vector machine"[MeSH Terms] OR ("support"[All Fields] AND "vector"[All Fields] AND "machine"[All Fields]) OR "support vector machine"[All Fields] OR "support vector machine"[MeSH Terms]	26 805

STRATÉGIE de RECHERCHE MEDLINE (Pubmed)

N° de la recherche	Demande	Filtres	Détails de la recherche	Résultats
19	#11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18		"artificial intelligence"[MeSH Terms] OR ("artificial"[All Fields] AND "intelligence"[All Fields]) OR "artificial intelligence"[All Fields] OR "artificial intelligence"[MeSH Terms] OR (((("neural"[All Fields] OR "neuralization"[All Fields] OR "neuralize"[All Fields] OR "neuralized"[All Fields] OR "neuralizes"[All Fields] OR "neuralizing"[All Fields] OR "neurally"[All Fields]) AND "network*" [All Fields]) OR "neural networks, computer"[MeSH Terms]) OR ("machine learning"[MeSH Terms] OR ("machine"[All Fields] AND "learning"[All Fields]) OR "machine learning"[All Fields] OR "machine learning"[MeSH Terms]) OR (("machine"[All Fields] OR "machine s"[All Fields] OR "machines"[All Fields] OR ("computability"[All Fields] OR "computable"[All Fields] OR "computating"[All Fields] OR "computation"[All Fields] OR "computational"[All Fields] OR "computations"[All Fields] OR "compute"[All Fields] OR "computed"[All Fields] OR "computer s"[All Fields] OR "computers"[MeSH Terms] OR "computers"[All Fields] OR "computer"[All Fields] OR "computes"[All Fields] OR "computing"[All Fields] OR "computational"[All Fields])) AND ("vision s"[All Fields] OR "vision, ocular"[MeSH Terms] OR ("vision"[All Fields] AND "ocular"[All Fields]) OR "ocular vision"[All Fields] OR "vision"[All Fields] OR "visions"[All Fields] OR "visioning"[All Fields])) OR (("unsupervised"[All Fields] OR ("organization and administration"[MeSH Terms] OR ("organization"[All Fields] AND "administration"[All Fields]) OR "organization and administration"[All Fields] OR "supervision"[All Fields] OR "supervise"[All Fields] OR "supervised"[All Fields] OR "supervises"[All Fields] OR "supervising"[All Fields] OR "supervisions"[All Fields]) OR ("weak*" [All Fields] AND ("organization and administration"[MeSH Terms] OR ("organization"[All Fields] AND "administration"[All Fields]) OR "organization and administration"[All Fields] OR "supervision"[All Fields] OR "supervise"[All Fields] OR "supervised"[All Fields] OR "supervises"[All Fields] OR "supervising"[All Fields] OR "supervisions"[All Fields])) OR (("limit"[All Fields] OR "limitation"[All Fields] OR "limitations"[All Fields] OR "limited"[All Fields] OR "limiting"[All Fields] OR "limits"[All Fields]) AND ("organization and administration"[MeSH Terms] OR ("organization"[All Fields] AND "administration"[All Fields]) OR "organization and administration"[All Fields] OR "supervision"[All Fields] OR "supervise"[All Fields] OR "supervised"[All Fields] OR "supervises"[All Fields] OR "supervising"[All Fields] OR "supervisions"[All Fields])) AND ("learning"[MeSH Terms] OR "learning"[All Fields] OR "learn"[All Fields] OR "learned"[All Fields] OR "learning s"[All Fields] OR "learnings"[All Fields] OR "learns"[All Fields] OR "network"[All Fields] OR "network s"[All Fields] OR "networked"[All Fields] OR "networker"[All Fields] OR "networkers"[All Fields] OR "networking"[All Fields] OR "networks"[All Fields]) OR ("algorithm s"[All Fields] OR "algorithmic"[All Fields] OR "algorithmically"[All Fields] OR "algorithmics"[All Fields] OR "algorithmization"[All Fields] OR "algorithms"[MeSH Terms] OR "algorithms"[All Fields] OR "algorithm"[All Fields])) OR "diagnosis, computer assisted"[MeSH Terms] OR ("deep learning"[MeSH Terms] OR ("deep"[All Fields] AND "learning"[All Fields]) OR "deep learning"[All Fields] OR "deep learning"[MeSH Terms]) OR ("support vector machine"[MeSH Terms] OR ("support"[All Fields] AND "vector"[All Fields] AND "machine"[All Fields]) OR "support vector machine"[All Fields] OR "support vector machine"[MeSH Terms])	599 682

STRATÉGIE de RECHERCHE MEDLINE (Pubmed)

N° recherche	Demande	Filtres	Détails de la recherche	Résultats
20	#3 AND #10 AND #19		(((("early"[All Fields] OR "earlier"[All Fields] OR "facilitat*"[All Fields] OR "aid"[All Fields] OR "augment*"[All Fields] OR "improv*"[All Fields]) AND ("detect*"[All Fields] OR "diagnos*"[All Fields] OR "recogni*"[All Fields] OR "identif*"[All Fields] OR "classif*"[All Fields] OR "decision*"[All Fields])) OR ("triage"[MeSH Terms] OR "triage"[All Fields] OR "trriages"[All Fields] OR "triaged"[All Fields] OR "triaging"[All Fields])) AND ("melanoma"[MeSH Terms] OR ("melanoma"[MeSH Terms] OR "melanoma"[All Fields] OR "melanomas"[All Fields] OR "melanoma s"[All Fields]) OR (("suspicious"[All Fields] OR "suspiciously"[All Fields] OR ("pigment"[All Fields] OR "pigmental"[All Fields] OR "pigmentation"[MeSH Terms] OR "pigmentation"[All Fields] OR "pigmentations"[All Fields] OR "pigmented"[All Fields] OR "pigmenting"[All Fields] OR "pigmentous"[All Fields] OR "pigments"[All Fields])) AND ("skin neoplasms"[MeSH Terms] OR ("skin"[All Fields] AND "neoplasms"[All Fields]) OR "skin neoplasms"[All Fields] OR ("skin"[All Fields] AND "tumor"[All Fields]) OR "skin tumor"[All Fields] OR ("skin"[All Fields] AND "tumour"[All Fields]) OR "skin tumour"[All Fields] OR (("skin"[MeSH Terms] OR "skin"[All Fields]) AND ("lesion"[All Fields] OR "lesion s"[All Fields] OR "lesional"[All Fields] OR "lesions"[All Fields]))) OR ("carcinoma, basal cell"[MeSH Terms] OR ("carcinoma"[All Fields] AND "basal"[All Fields] AND "cell"[All Fields]) OR "basal cell carcinoma"[All Fields] OR ("basal"[All Fields] AND "cell"[All Fields] AND "carcinoma"[All Fields]) OR "carcinoma, basal cell"[MeSH Terms]) OR ("carcinoma, squamous cell"[MeSH Terms] OR ("carcinoma"[All Fields] AND "squamous"[All Fields] AND "cell"[All Fields]) OR "squamous cell carcinoma"[All Fields] OR ("squamous"[All Fields] AND "cell"[All Fields] AND "carcinoma"[All Fields]) OR "carcinoma, squamous cell"[MeSH Terms]) OR ("skin neoplasms"[MeSH Terms] OR ("skin"[All Fields] AND "neoplasms"[All Fields]) OR "skin neoplasms"[All Fields] OR ("skin"[All Fields] AND "cancer"[All Fields]) OR "skin cancer"[All Fields] OR "skin neoplasm*"[MeSH Terms]) OR (("skin"[MeSH Terms] OR "skin"[All Fields]) AND "malignan*"[All Fields])) AND ("artificial intelligence"[MeSH Terms] OR ("artificial"[All Fields] AND "intelligence"[All Fields]) OR "artificial intelligence"[All Fields] OR "artificial intelligence"[MeSH Terms] OR (((("neural"[All Fields] OR "neuralization"[All Fields] OR "neuralize"[All Fields] OR "neuralized"[All Fields] OR "neuralizes"[All Fields] OR "neuralizing"[All Fields] OR "neurally"[All Fields]) AND "network*"[All Fields]) OR "neural networks, computer"[MeSH Terms]) OR ("machine learning"[MeSH Terms] OR ("machine"[All Fields] AND "learning"[All Fields]) OR "machine learning"[All Fields] OR "machine learning"[MeSH Terms]) OR (("machine"[All Fields] OR "machine s"[All Fields] OR "machines"[All Fields] OR ("computability"[All Fields] OR "computable"[All Fields] OR "computating"[All Fields] OR "computation"[All Fields] OR "computational"[All Fields] OR "computations"[All Fields] OR "compute"[All Fields] OR "computed"[All Fields] OR "computer s"[All Fields] OR "computers"[MeSH Terms] OR "computers"[All Fields] OR "computer"[All Fields] OR "computes"[All Fields] OR "computing"[All Fields] OR "computational"[All Fields])) AND ("vision s"[All Fields] OR "vision, ocular"[MeSH Terms] OR ("vision"[All Fields] AND "ocular"[All Fields]) OR "ocular vision"[All Fields] OR "vision"[All Fields] OR "visions"[All Fields] OR "visioning"[All Fields])) OR (("unsupervised"[All Fields] OR ("organization and administration"[MeSH Terms] OR ("organization"[All Fields] AND "administration"[All Fields]) OR "organization and administration"[All Fields] OR "supervision"[All Fields] OR "supervise"[All Fields] OR "supervised"[All Fields] OR "supervises"[All Fields] OR "supervising"[All Fields] OR "supervisions"[All Fields]) OR ("weak*"[All Fields] AND ("organization and administration"[MeSH Terms] OR ("organization"[All Fields] AND "administration"[All Fields]) OR "organization and administration"[All Fields] OR "supervision"[All Fields] OR "supervised"[All Fields] OR "supervises"[All Fields] OR "supervising"[All Fields] OR "supervisions"[All Fields]) OR "supervising"[All Fields] OR "supervisions"[All Fields])) OR (("limit"[All Fields] OR "limitation"[All Fields] OR "limitations"[All Fields] OR "limited"[All Fields] OR "limiting"[All Fields] OR "limits"[All Fields]) AND ("organization and administration"[MeSH Terms] OR ("organization"[All Fields] AND "administration"[All Fields]) OR "organization and administration"[All Fields] OR "supervision"[All Fields] OR "supervise"[All Fields] OR "supervised"[All Fields] OR "supervises"[All Fields] OR "supervising"[All Fields] OR "supervisions"[All Fields])))) AND ("learning"[MeSH Terms] OR "learning"[All Fields] OR "learn"[All Fields] OR "learned"[All Fields] OR "learning s"[All Fields] OR "learnings"[All Fields] OR "learns"[All Fields] OR ("network"[All Fields] OR "network s"[All Fields] OR "networked"[All Fields] OR "networker"[All Fields] OR "networkers"[All Fields] OR "networking"[All Fields] OR "networks"[All Fields]) OR ("algorithm s"[All Fields] OR "algorithmic"[All Fields] OR "algorithmically"[All Fields] OR "algorithms"[All Fields] OR "algorithmics"[All Fields] OR "algorithmization"[All Fields] OR "algorithms"[MeSH Terms] OR "algorithms"[All Fields] OR "algorithm"[All Fields])) OR "diagnosis, computer assisted"[MeSH Terms] OR ("deep learning"[MeSH Terms] OR ("deep"[All Fields] AND "learning"[All Fields]) OR "deep learning"[All Fields] OR "deep learning"[MeSH Terms]) OR ("support vector machine"[MeSH Terms] OR ("support"[All Fields] AND "vector"[All Fields] AND "machine"[All Fields]) OR "support vector machine"[All Fields] OR "support vector machine"[MeSH Terms]))	1 994

N° de la recherche	Demande	Filtres	Détails de la recherche	Résultats
21	#7 AND #14 AND #23	From 2021 – 2023	<p>((("early"[All Fields] OR "earlier"[All Fields] OR "facilitat*"[All Fields] OR "aid"[All Fields] OR "augment*"[All Fields] OR "improv*"[All Fields]) AND ("detect*"[All Fields] OR "diagnos*"[All Fields] OR "recogni*"[All Fields] OR "identif*"[All Fields] OR "classif*"[All Fields] OR "decision*"[All Fields])) OR ("triage"[MeSH Terms] OR "triage"[All Fields] OR "triages"[All Fields] OR "triaged"[All Fields] OR "triaging"[All Fields])) AND ("melanoma"[MeSH Terms] OR ("melanoma"[MeSH Terms] OR "melanoma"[All Fields] OR "melanomas"[All Fields] OR "melanoma s"[All Fields]) OR (("suspicious"[All Fields] OR "suspiciously"[All Fields] OR ("pigment"[All Fields] OR "pigmental"[All Fields] OR "pigmentation"[MeSH Terms] OR "pigmentation"[All Fields] OR "pigmentations"[All Fields] OR "pigmented"[All Fields] OR "pigmenting"[All Fields] OR "pigmentous"[All Fields] OR "pigments"[All Fields])) AND ("skin neoplasms"[MeSH Terms] OR ("skin"[All Fields] AND "neoplasms"[All Fields]) OR "skin neoplasms"[All Fields] OR ("skin"[All Fields] AND "tumor"[All Fields]) OR "skin tumor"[All Fields] OR ("skin"[All Fields] AND "tumour"[All Fields]) OR "skin tumour"[All Fields] OR (("skin"[MeSH Terms] OR "skin"[All Fields]) AND ("lesion s"[All Fields] OR "lesional"[All Fields] OR "lesions"[All Fields]))) OR ("carcinoma, basal cell"[MeSH Terms] OR ("carcinoma"[All Fields] AND "basal"[All Fields] AND "cell"[All Fields]) OR "basal cell carcinoma"[All Fields] OR ("basal"[All Fields] AND "cell"[All Fields] AND "carcinoma"[All Fields]) OR "carcinoma, basal cell"[MeSH Terms]) OR ("carcinoma, squamous cell"[MeSH Terms] OR ("carcinoma"[All Fields] AND "squamous"[All Fields] AND "cell"[All Fields]) OR "squamous cell carcinoma"[All Fields] OR ("squamous"[All Fields] AND "cell"[All Fields] AND "carcinoma"[All Fields]) OR "carcinoma, squamous cell"[MeSH Terms]) OR ("skin neoplasms"[MeSH Terms] OR ("skin"[All Fields] AND "neoplasms"[All Fields]) OR "skin neoplasms"[All Fields] OR ("skin"[All Fields] AND "cancer"[All Fields]) OR "skin cancer"[All Fields] OR ("skin neoplasm*"[MeSH Terms] OR ("skin"[MeSH Terms] OR "skin"[All Fields]) AND "malignan*"[All Fields])) AND ("artificial intelligence"[MeSH Terms] OR ("artificial"[All Fields] AND "intelligence"[All Fields]) OR "artificial intelligence"[All Fields] OR "artificial intelligence"[MeSH Terms] OR (((("neural"[All Fields] OR "neuralization"[All Fields] OR "neuralize"[All Fields] OR "neuralized"[All Fields] OR "neuralizes"[All Fields] OR "neuralizing"[All Fields] OR "neurally"[All Fields]) AND "network*"[All Fields]) OR "neural networks, computer"[MeSH Terms]) OR ("machine learning"[MeSH Terms] OR ("machine"[All Fields] AND "learning"[All Fields]) OR "machine learning"[All Fields] OR "machine learning"[MeSH Terms]) OR ("machine"[All Fields] OR "machine s"[All Fields] OR "machines"[All Fields] OR ("computability"[All Fields] OR "computable"[All Fields] OR "computating"[All Fields] OR "computation"[All Fields] OR "computational"[All Fields] OR "computations"[All Fields] OR "compute"[All Fields] OR "computed"[All Fields] OR "computer s"[All Fields] OR "computers"[MeSH Terms] OR "computers"[All Fields] OR "computer"[All Fields] OR "computes"[All Fields] OR "computing"[All Fields] OR "computional"[All Fields])) AND ("vision s"[All Fields] OR "vision, ocular"[MeSH Terms] OR ("vision"[All Fields] AND "ocular"[All Fields]) OR "ocular vision"[All Fields] OR "vision"[All Fields] OR "visions"[All Fields] OR "visioning"[All Fields])) OR (("unsupervised"[All Fields] OR ("organization and administration"[MeSH Terms] OR ("organization"[All Fields] AND "administration"[All Fields]) OR "organization and administration"[All Fields] OR "supervision"[All Fields] OR "supervise"[All Fields] OR "supervised"[All Fields] OR "supervises"[All Fields] OR "supervising"[All Fields] OR "supervisions"[All Fields]) OR ("weak*"[All Fields] AND ("organization and administration"[MeSH Terms] OR ("organization"[All Fields] AND "administration"[All Fields]) OR "organization and administration"[All Fields] OR "supervision"[All Fields] OR "supervise"[All Fields] OR "supervised"[All Fields] OR "supervises"[All Fields] OR "supervising"[All Fields] OR "supervisions"[All Fields])) OR ("limit"[All Fields] OR "limitation"[All Fields] OR "limitations"[All Fields] OR "limited"[All Fields] OR "limiting"[All Fields] OR "limits"[All Fields]) AND ("organization and administration"[MeSH Terms] OR ("organization"[All Fields] AND "administration"[All Fields]) OR "organization and administration"[All Fields] OR "supervision"[All Fields] OR "supervise"[All Fields] OR "supervised"[All Fields] OR "supervises"[All Fields] OR "supervising"[All Fields] OR "supervisions"[All Fields])) AND ("learning"[MeSH Terms] OR "learning"[All Fields] OR "learn"[All Fields] OR "learned"[All Fields] OR "learning s"[All Fields] OR "learnings"[All Fields] OR "learns"[All Fields] OR ("network"[All Fields] OR "network s"[All Fields] OR "networked"[All Fields] OR "networker"[All Fields] OR "networkers"[All Fields] OR "networking"[All Fields] OR "networks"[All Fields]) OR ("algorithm s"[All Fields] OR "algorithmic"[All Fields] OR "algorithmically"[All Fields] OR "algorithmics"[All Fields] OR "algorithmization"[All Fields] OR "algorithms"[MeSH Terms] OR "algorithms"[All Fields] OR "algorithm"[All Fields])) OR "diagnosis, computer assisted"[MeSH Terms] OR ("deep learning"[MeSH Terms] OR ("deep"[All Fields] AND "learning"[All Fields]) OR "deep learning"[All Fields] OR "deep learning"[MeSH Terms]) OR ("support vector machine"[MeSH Terms] OR ("support"[All Fields] AND "vector"[All Fields] AND "machine"[All Fields]) OR "support vector machine"[All Fields] OR "support vector machine"[MeSH Terms])) AND (2021:2023[pdat])</p>	968

VII-2 ANNEXE 2 : Liste des articles exclus

TITRE	AUTEUR	REVUE	ANNÉE	EXCLUSION
A Secure Framework toward IoMT-Assisted Data Collection, Modeling, and Classification for Intelligent Dermatology Healthcare Services	MK. Islam, C. Kaushal, A. Amin, A.D Algarni, N. Alturki, N.F Soliman, R.F Mansour	Contrast Media and Molecular Imaging	2022	Population
A smartphone-based application for an early skin disease prognosis: Towards a lean healthcare system via computer-based vision	M. Shahin, F.F. Chen, Ali Hosseinzadeh, H.K. Koodiani, A. Shahin, O.A. Nafi	Advanced Engineering Informatics	2023	Population
AI-assisted mole detection for online dermatology triage in telemedicine settings	D. Das, E. Ergin, B. Morel, M. Noga, D. Emery, K. Punithakumar	Informatics in Medecine Unlocked	2023	Population
An Automated Deep Learning based Ensemble Approach for Malignant Melanoma Detection using Dermoscopy Images	K. Safdar, S. Akbar, S. Gull	International Conference on Frontiers of Information Technology	2021	Population
An interpretable CNN-based CAD system for skin lesion diagnosis.	J. López-Labraca, I. González-Díaz, F. Díaz-de-María, A. Fueyo-Casado	Artificial Intelligence in Medecine	2022	Population
An Interpretable Skin Cancer Classification Using Optimized Convolutional Neural Network for a Smart Healthcare System	K. Mridha, M. Uddin, J. Shin, S. Khadka, M.F. Mridha	IEEE Access	2023	Population
Artificial intelligence support for skin lesion triage in primary care and dermatology	H. Thompson, A. Oakley, M. B Jameson, A. Bowling	Australasian Journal of Dermatology	2021	Hors article
Assessment of melanoma thickness based on dermoscopy images: an open, web-based, international, diagnostic study	S. Polesie, M. Gillstedt, H. Kittler, C. Rinner, P. Tschandl, J. Paoli	Journal of the European Academy of Dermatology and Venereology	2022	Hors sujet d'étude
Automatic identification of benign pigmented skin lesions from clinical images using deep convolutional neural network	H. Ding, E. Zhang, F. Fang, X. Liu, H. Zheng, H. Yang, Y. Ge, Y. Yang, T. Lin	BMC Biotechnology	2022	Hors sujet d'étude
Clinical performance of an artificial intelligence based medical device deployed within an urgent suspected skin cancer pathway	R. Jenkins, C.F. Brewer, D. Kalsi, D. Mullarkey	British Journal of Dermatology	2023	Hors article
Community and Locality Imaging Centres: Innovation in the skin cancer care pathway	S. Muthiah, A. Coon, A. Bryden, G. Chin, C. Fleming, C. Morton	British Journal of Dermatology	2023	Hors article
DermlA: Machine Learning to Improve Skin Cancer Screening.	E. Shoen	Journal of Digital Imaging	2021	Facteur d'impact
Development and Clinical Evaluation of an Artificial Intelligence Support Tool for Improving Telemedicine Photo Quality.	K. Vodrahalli, J. Ko, A.S. Chiou, R. Novoa, A. Abid, M. Phung, K. Yekrang, P. Petrone, J. Zou, R. Daneshjou	JAMA Dermatology	2023	Hors sujet d'étude
Development and Narrow Validation of Computer Vision Approach to Facilitate Assessment of Change in Pigmented Cutaneous Lesions.	W.F. Maguire, P.H. Haley, C.M. Dietz, M. Hoffelder, C.S Brandt, R. Joyce, G. Fitzgerald, C. Minnier, C. Sander, L.K. Ferris, G. Paragh, J. Arbesman, H. Wang, K.J. Mitchell, E.K. Hughes, J.M. Kirkwood	JID Innovations	2023	Population
Discrimination Between Invasive and In Situ Melanomas Using Clinical Close-Up Images and a De Novo Convolutional Neural Network.	S. Polesie, M. Gillstedt, G. Ahlgren, H. Ceder, J. Dahlén Gyllencreutz, J. Fougelberg, E. Johansson Backman, J. Pakka, O. Zaar, J. Paoli	Frontiers in Medecine	2021	Hors sujet d'étude
Evaluation of a smartphone application for diagnosis of skin diseases.	M. Mikołajczyk, S. Patrzyk, M. Nieniewski, A. Woźniacka	Postepy dermatologii i alergologii	2021	Hors sujet d'étude
Improving Teledermatology Referral with Edge-AI: Mobile App to Foster Skin Lesion Imaging Standardization	M.J.M. Vasconcelos, D. Moreira, P. Alves, R. Graça, R. Franco, L. Rosado	Biomedical Engineering Systems and Technologies	2023	Hors article
Incorporating clinical knowledge with constrained classifier chain into a multimodal deep network for melanoma detection.	Y. Wang, J. Cai, D.C. Louie, Z.J. Wang, T. K. Lee	Computers in Biology and Medecine	2021	Population

TITRE	AUTEUR	REVUE	ANNÉE	EXCLUSION
Interpretable Skin Cancer Classification based on Incremental Domain Knowledge Learning.	<u>E. Rezk, M. Eltorki, W. El-Dakhakhni</u>	Journal of Healthcare Informatics Research	2023	Population
Melanoma Detection Using XGB Classifier Combined with Feature Extraction and K-Means SMOTE Techniques.	C.C. Chang, Y.Z. Li, H.C. Wu, M.H. Tseng	Diagnostics (Basel)	2022	Population
MIoT Based Skin Cancer Detection Using Bregman Recurrent Deep Learning	N.R. Sivakumar, S. A. Ghorashi, F.K. Karim, E. Alabdulkreem, A. Al-Rasheed	Computers, Materials & Continua	2022	Population
Prepare to succeed: real-world image datasets for artificial intelligence in skin cancer triage	G.X.M. Chin, T. Suveges, J. Carse, S. Butt, S. Muthiah, C. Morton, E. Trucco, C. Proby, S. McKenna, C. Fleming	British Journal of Dermatology	2022	Hors article
Reproducible Naevus Counts Using 3D Total Body Photography and Convolutional Neural Networks.	B. Betz-Stablein, B. D'Alessandro, U. Koh, E. Plasmeijer, M. Janda, S.W. Menzies, R. Hofmann-Wellenhof, A.C. Green, H.P. Soyer	Dermatology (Basel)	2022	Hors sujet d'étude
Skin Cancer Detection Based on Deep Learning.	R. Ahmadi Mehr, A. Ameri	Journal of biomedical physics	2022	Population
Skin cancer risk self-assessment using AI as a mass screening tool	A.O. Ukharov, I.L. Shlivko, I.A. Klemenova O.E. Garanina, K.A. Uskova, A .M. Mironycheva, Y.L. Stepanova	Informatics in Medecine Unlocked	2023	Facteur d'impact
Skin lesion image classification method based on extension theory and deep learning	X.F Bian, H.W. Pan, K.J. Zhang, P.Y. Li, J.B. Li, C.L. Chen	Multimedia Tools and Applications	2022	Population
Using artificial intelligence to triage skin cancer referrals: Outcomes from a pilot study	K. Abu Baker, E. Roberts, K. Harman, D. Mullarkey, D. Kalsi	British Journal of Dermatology	2023	Hors article
Validation of a Market-Approved Artificial Intelligence Mobile Health App for Skin Cancer Screening: A Prospective Multicenter Diagnostic Accuracy Study.	T. Sangers, S. Reeder, S. van der Vet, S. Jhingoer, A. Mooyaart, D.M. Siegel, T. Nijsten, M. Wakkee	Dermatology (Basel)	2022	Schéma d'étude

VII- 3 ANNEXE 3 : Évaluation de la qualité et de la transparence des articles selon les échelles STROBE, TREND et CONSORT

Note : Pour toutes les échelles, les items sont notés 0 (non rapporté), 1 (rapporté) ou * (non concerné).

Échelle STROBE :

Section / Sujet	Item	Recommandations	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
TITRE ET RÉSUMÉ																							
	1a	Indiquer dans le titre ou dans le résumé le type d'étude réalisée en termes couramment utilisés.	0	0	0	1	1	0	0	1	0	1	1	0	1	0	0	0	1	0	1	0	0
	1b	Fournir dans le résumé une information synthétique et objective sur ce qui a été fait et sur les résultats.	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
INTRODUCTION																							
Contexte / Justification	2	Expliquer le contexte scientifique et la légitimité de l'étude en question.	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Objectifs	3	Citer les objectifs spécifiques, y compris toutes les hypothèses a priori.	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
MÉTHODES																							
Conception de l'étude	4	Présenter les éléments clés de la conception de l'étude en tout début de document.	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1
Contexte / Justification	5	Décrire le contexte, les lieux et les dates pertinentes, y compris les périodes de recrutement, d'exposition, de suivi et de recueil de données.	1	1	0	1	1	1	1	1	0	1	1	1	1	1	0	0	1	1	1	0	1
Population	6a	Étude de cohorte – Indiquer les critères d'éligibilité, les sources et méthodes de sélection des sujets. Décrire les méthodes de suivi. Étude cas-témoin – Indiquer les critères d'éligibilité et les sources et méthodes pour identifier les cas et sélectionner les témoins. Justifier le choix des cas et des témoins. Étude transversale – Indiquer les critères d'éligibilité, les sources et méthodes de sélection des participants.	1	0	1	*	1	1	*	1	0	1	1	1	0	*	*	0	1	1	1	0	0
	6b	Étude de cohorte – Pour les études appariées, indiquer les critères d'appariement et le nombre de sujets exposés et non exposés. Étude cas-témoin – Pour les études appariées, indiquer les critères d'appariement et le nombre de témoins par cas.	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Variables	7	Définir clairement tous les critères de résultats, les expositions, les facteurs de prédiction, les facteurs de confusion potentiels et les facteurs d'influence. Indiquer les critères diagnostiques, le cas échéant.	0	*	1	1	0	1	*	1	1	0	*	0	1	0	0	0	1	0	1	1	1
Sources de données / mesures	8	Pour chaque variable d'intérêt, indiquer les sources de données et les détails des méthodes d'évaluation (mesures). Décrire la comparabilité des méthodes d'évaluation s'il y a plus d'un groupe.	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Section / Sujet	Item	Recommandations	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
Biais	9	Décrire toutes les mesures prises pour éviter les sources potentielles de biais.	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	
Taille de l'étude	10	Expliquer comment a été déterminé le nombre de sujets à inclure.	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	
Variables quantitatives	11	Expliquer comment les variables quantitatives ont été traitées dans les analyses. Le cas échéant, décrire quels regroupements ont été effectués et pourquoi.	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1
Analyses statistiques	12a	Décrire toutes les analyses statistiques, y compris celles utilisées pour contrôler les facteurs de confusion.	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	12b	Décrire toutes les méthodes utilisées pour examiner les sous-groupes et les interactions.	1	*	*	*	1	0	1	1	1	0	*	0	*	0	0	*	*	*	*	*	*	
	12c	Expliquer comment les données manquantes ont été traitées.	0	*	*	0	*	0	*	*	*	0	*	0	*	0	0	*	0	*	*	0	*	
	12d	Étude de cohorte – Le cas échéant, expliquer comment les perdus de vue ont été traités. Étude cas-témoin – Le cas échéant, expliquer comment l'appariement des cas et des témoins a été réalisé. Étude transversale – Le cas échéant, décrire les méthodes d'analyse qui tiennent compte de la stratégie d'échantillonnage.	*	*	1	*	0	*	*	*	*	*	*	1	0	*	*	0	*	0	0	*	0	
	12e	Décrire toutes les analyses de sensibilité.	1	1	0	0	0	0	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	
RÉSULTATS																								
Population	13a	Rapporter le nombre d'individus à chaque étape de l'étude – par exemple : potentiellement éligibles, examinés pour l'éligibilité, confirmés éligibles, inclus dans l'étude, complètement suivis et analysés.	0	1	0	*	0	1	1	1	0	1	1	0	1	*	0	1	0	1	1	1	1	
	13b	Indiquer les raisons de non-participation à chaque étape.	*	*	*	*	0	0	*	*	*	*	*	0	*	*	*	*	0	*	0	0	*	
	13c	Envisager l'utilisation d'un diagramme de flux.	1	1	0	0	0	0	1	0	0	1	1	1	0	1	1	0	0	1	0	0	0	
Données descriptives	14a	Indiquer les caractéristiques de la population étudiée (par exemple : démographiques, cliniques, sociales) et les informations sur les expositions et les facteurs de confusion potentiels.	1	0	0	*	1	1	0	1		1	0	0	1	0	0	0	1	1	1	1	0	
	14b	Indiquer le nombre de sujets inclus avec des données manquantes pour chaque variable d'intérêt.	*	1	1	*	0	0	*	1	1	1	1	*	1	1	0	1	0	*	1	1	1	
	14c	Étude de cohorte – Résumer la période de suivi (par exemple : nombre moyen et total).	*	*	*	*	0	*	*	0	*	*	1	*	*	*	*	0	1	1	1	*	*	
Données obtenues	15	Étude de cohorte – Rapporter le nombre d'évènements survenus ou les indicateurs mesurés au cours du temps. Étude cas-témoin – Reporter le nombre de sujets pour chaque catégorie d'exposition ou les indicateurs du niveau d'exposition mesurés. Étude transversale – Reporter le nombre d'évènements survenus ou les indicateurs mesurés.	1	*	1	*	1	1	1	1	*	1	*	1	1	*	*	*	0	0	0	1	1	

Section / Sujet	Item	Recommandations	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Principaux résultats	16a	Indiquer les estimations non ajustées et, le cas échéant, les estimations après ajustement sur les facteurs de confusion avec leur précision (ex : intervalle de confiance de 95 %). Expliciter quels facteurs de confusion ont été pris en compte et pourquoi ils ont été inclus.	1	1	0	1	1	1	1	0	0	0	0	0	1	0	0	1	1	0	*	0	0
	16b	Indiquer les valeurs bornes des intervalles lorsque les variables continues ont été catégorisées.	*	*	*	0	1	0	1	1	0	1	1	0	*	1	1	*	*	0	1	*	1
	16c	Selon les situations, traduire les estimations de risque relatif en risque absolu sur une période (cliniquement) interprétable.	*	*	*	*	*	0	*	*	*	*	*	*	0	0	*	*	*	0	*	*	*
Autres analyses	17	Mentionner les autres analyses réalisées – Ex : analyses de sous-groupes, recherche d'interactions et analyses de sensibilité.	*	1	0	1	1	*	*	1	1	1	1	1	1	1	1	1	0	1	*	0	1
DISCUSSION																							
Résultats clés	18	Résumer les principaux résultats en se référant aux objectifs de l'étude.	1	1	1	1	1	1	1	1	1	0	1	1	1	0	0	1	0	1	1	1	1
Limitations	19	Discuter les limites de l'étude, en tenant compte des sources de biais potentiels ou d'imprécisions. Discuter du sens et de l'importance de tout biais potentiel.	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Interprétation	20	Donner une interprétation générale prudente des résultats compte tenu des objectifs, des limites de l'étude, de la multiplicité des analyses, des résultats d'études similaires et de tout autre élément pertinent.	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1
Généralisabilité	21	Discuter la « généralisabilité » (validité externe) des résultats de l'étude.	1	1	1	1	1	1	0	1		1	1	0	0	1	0	0	0	1	0	1	1
AUTRE INFORMATION																							
Financement	22	Indiquer la source de financement et le rôle des financeurs pour l'étude rapportée, le cas échéant, pour l'étude originale sur laquelle s'appuie l'article présenté.	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1
TOTAL / 34			17	19	16	17	21	19	20	24	17	20	23	17	21	15	14	15	18	20	21	16	20

Échelle TREND :

Section / Sujet	Item	Description	22	23	24
TITRE ET RÉSUMÉ					
	1a	Description de la méthode utilisée pour assigner ou distribuer les participants ou autres unités (groupes ou communautés) aux différentes interventions	0	0	0
	1b	Résumé structuré du plan d'essai, méthodes, résultats et conclusions	1	1	1
	1c	Informations sur la population cible ou l'échantillon d'étude	1	1	1
INTRODUCTION					
	2a	Contexte scientifique et explication du bien fondé	1	1	1
	2b	Théories utilisées dans la conception des interventions comportementales	*	*	*
MÉTHODES					
Participants	3a	Critères d'éligibilité pour les participants, y compris les critères à différents niveaux du plan de recrutement/échantillonnage (par ex. villes, cliniques, sujets)	1	1	1
	3b	Méthode de recrutement (comment ? référence par médecin, auto-sélection...), y compris la méthode d'échantillonnage si un plan d'échantillonnage systématique a été mis en œuvre	1	1	1
	3c	Cadre de recrutement (où les participants sont-ils recrutés ? cliniques ...)	1	1	1
	3d	Cadres et lieux dans lesquels les données ont été collectées (laboratoires, cliniques ...)	1	1	1
Interventions		<i>Détails des interventions prévues pour chaque condition d'étude, comment et quand ont-elles été effectivement administrées, en particulier :</i>			
	4a	Contenu : description précise des éléments ou des composantes spécifiques de l'intervention fournis aux participants	1	1	1
	4b	Méthode de livraison : comment le contenu a-t-il été délivré ?	1	1	1
	4c	Unité de livraison : l'intervention a-t-elle été administrée de manière individuelle ou collective, comment les participants ont-ils été regroupés le cas échéant ?	1	1	0
	4d	Prestataire : qui a délivré l'intervention ?	1	1	1
	4e	Cadre : où l'intervention a-t-elle été délivrée ?	1	1	1
	4f	Quantité et durée d'exposition : combien de sessions ou d'épisodes étaient prévus ? Combien de temps étaient-ils censés durer ?	0	0	0
	4g	Période : Quelle durée était prévue pour l'administration de l'intervention à chaque participant/groupe ?	1	0	0
	4h	Activités pour augmenter la conformité ou l'adhésion aux protocoles de l'étude (incitations financières, support logistique, suivi personnalisé...)	0	0	0
Objectifs	5	Description des objectifs et hypothèses spécifiques	1	1	1
Résultats	6a	Mesures de résultat primaires et secondaires clairement définies	1	1	1
	6b	Méthodes utilisées pour collecter les données et toute méthode utilisée pour améliorer la qualité des mesures	1	1	1
	6c	Informations sur les instruments validés tels que les propriétés psychométriques et biométriques	0	0	0
Taille de l'échantillon	7	Comment la taille de l'échantillon a-t-elle été déterminée et, le cas échéant, explication de toute analyse intermédiaire et des règles d'arrêt	0	0	0
Attribution	8a	Unité d'attribution (l'unité étant assignée à la condition d'étude, par ex. individu, groupe, communauté)	1	1	1
	8b	Méthode utilisée pour assigner les unités aux conditions d'étude, y compris les détails de toute restriction (par ex., blocage, stratification, minimisation)	0	0	0
	8c	Inclusion des aspects mis en œuvre pour minimiser les biais potentiels dus à la non-randomisation (par ex., appariement)	0	0	0

Section / Sujet	Item	Description	22	23	24
Procédure d'aveuglement	9	Indication permettant de savoir si les participants, ceux qui administrent les interventions et ceux qui évaluent les résultats étaient en aveugle ; si oui, déclaration sur la manière dont l'aveuglement a été réalisé et évalué	1	0	1
Unité d'analyse	10a	Description de la plus petite unité analysée pour évaluer les effets de l'intervention (par ex., individu, groupe, communauté)	1	1	1
	10b	Si l'unité d'analyse diffère de l'unité d'attribution, la méthode analytique utilisée pour en tenir compte (par ex., ajustement des estimations d'erreur standard par l'effet de conception ou utilisation d'une analyse multiniveau)	*	*	0
Méthode statistique	11a	Méthodes statistiques utilisées pour comparer les groupes d'étude pour les résultats primaires, y compris les méthodes complexes pour les données corrélées	0	1	1
	11b	Méthodes statistiques utilisées pour des analyses supplémentaires, telles que les analyses de sous-groupes et les analyses ajustées	0	0	1
	11c	Méthodes pour imputer les données manquantes, si utilisées	0	0	0
	11d	Logiciel ou programmes statistiques utilisés	1	1	1
RÉSULTATS					
Flux des participants	12a	Flux des participants à travers chaque étape de l'étude : recrutement, attribution, allocation et exposition à l'intervention, suivi, analyse (un diagramme est fortement recommandé)	1	1	1
	12b	Recrutement : le nombre de participants dépistés pour l'éligibilité, jugés éligibles ou non, ayant refusés d'être inscrits et inscrits dans l'étude	1	0	0
	12c	Attribution : le nombre de participants assignés à une condition d'étude	*	1	1
	12d	Allocation et exposition à l'intervention : le nombre de participants ayant reçu chaque intervention	1	1	1
	12e	Suivi : le nombre de participants ayant complété le suivi ou non (par ex. perdus de vue), par condition d'étude	1	0	0
	12f	Analyse : le nombre de participants inclus ou exclus de l'analyse principale, par condition d'étude	0	1	0
	12g	Description des écarts par rapport au protocole de l'étude tel que prévu, ainsi que les raisons	0	0	0
Recrutement	13	Dates définissant les périodes de recrutement et de suivi	1	1	1
Données initiales	14a	Caractéristiques démographiques et cliniques de base (avant l'intervention) des participants dans chaque condition d'étude	1	1	1
	14b	Caractéristiques initiales pour chaque condition d'étude pertinentes pour la recherche en prévention des maladies (documenter et comparer les informations initiales directement liées au risque de développer la maladie étudiée)	*	*	*
	14c	Comparaisons de base entre ceux perdus de vue et ceux retenus, globalement et par condition d'étude	0	0	0
	14d	Comparaison entre la population de l'étude au départ et la population cible d'intérêt	0	0	0
Equivalence initiale	15	Données sur l'équivalence des groupes d'étude au départ et méthodes statistiques utilisées pour contrôler les différences initiales	*	0	1
Participants analysés	16a	Nombre de participants (dénominateur) inclus dans chaque analyse pour chaque condition d'étude, en particulier lorsque les dénominateurs changent pour différents résultats ; déclaration des résultats en nombres absolus lorsque cela est possible	0	1	*
	16b	Indication permettant de savoir si la stratégie d'analyse était "intention de traiter" ou, sinon, description de la manière dont les non-adhérents ont été traités dans les analyses	0	0	0
Résultats et estimation	17a	Pour chaque résultat primaire et secondaire, un résumé des résultats pour chaque condition d'étude, l'estimation de la taille de l'effet et un intervalle de confiance pour indiquer la précision	1	1	1
	17b	Inclusion des résultats nuls et négatifs	1	1	1
	17c	Inclusion des résultats des tests des voies causales préspecifiées par lesquelles l'intervention était censée opérer, le cas échéant	*	*	*
Analyses auxiliaires	18	Résumé des autres analyses effectuées, y compris les analyses de sous-groupes ou restreintes, indiquant celles qui sont préspecifiées ou exploratoires	1	0	1

Section / Sujet	Item	Description	22	23	24
Evènements indésirables	19	Résumé de tous les événements indésirables importants ou effets inattendus dans chaque condition d'étude (y compris les mesures de résumé, les estimations de la taille de l'effet et les intervalles de confiance)	0	0	0
DISCUSSION					
Interprétation	20a	Interprétation des résultats, en tenant compte des hypothèses de l'étude, des sources potentielles de biais, de l'imprécision des mesures, des analyses multiplicatives, et d'autres limites ou faiblesses de l'étude	1	1	1
	20b	Discussion des résultats de l'étude : sont-ils en ligne avec le mécanisme théorique de l'intervention ou existe-t-il des mécanismes ou explications alternatives ?	1	1	1
	20c	Discussion du succès et des obstacles à la mise en œuvre de l'intervention, de la fidélité de la mise en œuvre	1	0	0
	20d	Discussion des implications pour la recherche, les programmes ou les politiques	1	1	1
Généralisabilité	21	Généralisabilité (validité externe) des résultats de l'essai, en tenant compte de la population étudiée, des caractéristiques de l'intervention, de la durée du suivi, des incitations, des taux de conformité, des sites/cadres spécifiques impliqués dans l'étude, et d'autres questions contextuelles	1	1	1
Preuve globale	22	Interprétation générale des résultats dans le contexte des preuves actuelles et de la théorie actuelle	1	1	1
TOTAL / 58			35	33	34

Échelle CONSORT :

Section / Sujet	Item	Description	25
TITRE ET RÉSUMÉ			
	1a	Identification en tant qu'"essai randomisé" dans le titre.	1
	1b	Résumé structuré du plan d'essai, méthodes, résultats et conclusions (pour une aide spécifique voir CONSORT pour Abstract).	1
INTRODUCTION			
	2a	Contexte scientifique et explication du bien fondé.	1
	2b	Objectifs spécifiques et hypothèses.	1
MÉTHODES			
Plan de l'essai	3a	Description du plan de l'essai (tel que : groupes parallèles, plan factoriel) en incluant les ratios d'allocation.	1
	3b	Changements importants de méthode après le début de l'essai (tel que les critères d'éligibilité), en justifiant.	*
Participants	4a	Critères d'éligibilité des participants.	1
	4b	Structures et lieux de recueil des données.	1
Interventions	5	Interventions pour chaque groupe avec suffisamment de détails pour pouvoir reproduire l'étude, incluant comment et quand elles ont été véritablement conduites.	1
Critères de jugement	6a	Critères "a priori" de jugement principal et secondaires entièrement définis, en incluant comment et quand ils ont été évalués.	1
	6b	Changement quelconque de critères de jugement après le début de l'essai, en justifiant.	*
Taille de l'échantillon	7a	Comment la taille de l'échantillon a-t-elle été déterminée ?	0
	7b	Lorsque cela est applicable, explication des analyses intermédiaires et des règles d'arrêt.	1
RANDOMISATION			
Production de la séquence	8a	Méthode utilisée pour générer la séquence d'allocation par tirage au sort.	1
	8b	Type de randomisation, incluant les détails relatifs à une méthode de restriction (comme par ex : par blocs, avec leurs tailles).	0
	9	Mécanisme utilisé de mise en œuvre de la séquence d'allocation randomisée (par ex : utilisation d'enveloppes numérotées séquentiellement), décrivant chaque mesure prise pour masquer l'allocation jusqu'à l'assignation des interventions.	0
Mise en œuvre	10	Qui a généré la séquence d'allocation, qui a enrôlé les participants et qui a assigné les participants à leur groupe ?	0
Aveugle	11a	Au cas où, décrire qui a été en aveugle après l'assignation des interventions (par ex les participants, les administrateurs de traitement, les évaluateurs de résultats) et précisions sur la conduite du protocole en double aveugle.	*
	11b	Si approprié, description de la similitude des interventions.	1
Méthode statistique	12a	Méthodes statistiques utilisées pour comparer les groupes au regard des critères de jugement principal et secondaires.	1
	12b	Méthodes utilisées pour des analyses supplémentaires, telles que des analyses de sous-groupes ou des analyses ajustées.	1
RÉSULTATS			
Flux des participants	13a	Pour chaque groupe : nombre de participants qui ont été assignés par tirage au sort ayant reçu le traitement qui leur était destiné, et qui ont été analysés pour le critère de jugement principal.	1
	13a	Pour chaque groupe : abandons et exclusions après la randomisation, en justifiant.	1

Section / Sujet	Item	Description	25
Recrutement	14a	Dates définissant les périodes de recrutement et de suivi.	1
	14b	Pourquoi l'essai a-t-il pris fin ou a t'il été interrompu ?	*
Données initiales	15	Une table décrivant les caractéristiques initiales démographiques et cliniques de chaque groupe.	1
Effectifs analysés	16	Nombre de participants (dénominateur) inclus dans chaque analyse en précisant si elle a été effectuée avec les groupes d'origine.	1
Critères de jugement et estimations	17a	Pour chaque critère de jugement principal et secondaire : indiquer les résultats pour chaque groupe et la taille estimée de l'effet ainsi que son degré de précision (ex : intervalle de confiance à 95%).	1
	17b	Pour les variables binaires, une présentation de la taille de l'effet en valeurs absolues et relatives est recommandée.	*
Analyses accessoires	18	Résultats de toute analyse supplémentaire réalisée, en incluant les analyses en sous-groupes et les analyses ajustées et en distinguant les analyses spécifiées à priori des analyses exploratoires.	*
Risques	19	Tous les risques importants ou effets secondaires inattendus dans chaque groupe (pour un conseil détaillé voir CONSORT for Harms).	1
DISCUSSION			
Limitation	20	Limitations de l'essai, en tenant compte des sources de biais potentiels ou des imprécisions et, au cas où, en tenant compte de la multiplicité des analyses.	1
Généralisabilité	21	Généralisabilité (validité externe, applicabilité) des résultats de l'essai.	1
Interprétation	22	Interprétations/conclusions cohérentes avec les résultats, en tenant compte du ratio bénéfiques/risques et de possibles autres faits pertinents.	1
INFORMATIONS SUPPLÉMENTAIRES			
Enregistrement	23	Numéro d'enregistrement de l'essai en précisant le registre utilisé.	1
Protocole	24	A quel endroit, si possible, le protocole complet de l'essai peut-il être consulté ?	1
Financement	25	Sources de financement et autres ressources (ex : fourniture de médicaments), rôle des donateurs.	0
TOTAL / 37			26

VII-4 ANNEXE 4 : Évaluation de la méthodologie des articles selon les échelles QUADAS-2, MINORS et Jadad

Échelle QUADAS-2 :

Section / Sujet	Item	Description	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
SÉLECTION DES PATIENTS																							
Limitation du risque de biais	1a	Le recrutement des participants a-t-il été consécutif (tous les patients remplissant les critères d'inclusion sont inclus les uns après les autres, sans omission) ou aléatoire (sélection des participants au hasard parmi ceux qui remplissent les critères d'inclusion en assurant que chaque patient éligible a une chance égale d'être inclus dans l'étude) ?	0	1	0	0	0	1	0	0	0	1	1	0	1	0	0	0	1	0	0	0	0
	1b	Un schéma d'étude type cas-témoin a-t-il été évité ?	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	1c	L'étude a-t-elle évité des exclusions inappropriées ?	1	1	*	0	1	0	0	1	0	1	1	1	1	1	1	0	0	1	1	1	0
Applicabilité	2a	Y a-t-il un risque que les patients sélectionnés ne correspondent pas à la question de recherche ? (Les patients inclus sont-ils représentatifs de ceux pour lesquels le test est utilisé en pratique clinique) ? si oui 0 point, si non 1 point	0	0	0	1	0	0	0	0	0	1	0	0	0	0	1	1	0	0	1	1	0
TEST EVALUÉ																							
Limitation du risque de biais	3a	Les résultats du test évalué ont-ils été interprétés sans avoir connaissance des résultats du test de référence ?	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	3b	Si un seuil de décision a été utilisé, a-t-il été spécifié ?	*	*	*	1	1	0	*	1	1	1	*	1	1	1	*	1	*	1	1	*	*
Applicabilité	4a	Y a-t-il une préoccupation quant au fait que le test évalué, dans sa réalisation, interprétation, diffère de la question de recherche ? si oui 0 point, si non 1 point	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEST DE RÉFÉRENCE																							
Limitation du risque de biais	5a	Le test de référence est-il susceptible de correctement classer la condition recherchée (« gold standard ») ?	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
	5b	Les résultats du test de référence ont-ils été interprétés sans avoir connaissance des résultats du test à l'étude ?	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
Applicabilité	6a	Y a-t-il une préoccupation quant au fait que la pathologie telle qu'elle a été diagnostiquée par le test de référence diffère de celle de la question de recherche ? si oui 0 point, si non 1 point	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
DÉROULEMENT ET TEMPORALITÉ																							
Limitation du risque de biais	7a	L'intervalle de temps entre la réalisation du test évalué et le test de référence est-il approprié ?	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1
	7b	Tous les patients ont-ils reçu le test de référence ?	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
	7c	Tous les patients ont-ils reçu le même test de référence ?	0	0	1	0	0	0	1	0	0	0	1	0	1	0	0	0	1	0	1	0	0

Section / Sujet	Item	Description	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
	7d	Tous les patients inclus dans l'étude ont-ils été comptabilisés dans les résultats ?	1	1	1	1	1	1	0	1	0	0	1	0	1	0	0	1	1	1	1	1	0
TOTAL / 14			10	11	9	11	10	10	9	10	9	12	11	9	13	10	9	10	12	11	13	9	8

Échelle MINORS :

Section / Sujet	Item	Description	22	23	24
Objectif clairement annoncé	1	La question abordée doit être précise et pertinente à la lumière de la littérature disponible.	2	2	2
Inclusion de patients consécutifs	2	Tous les patients potentiellement éligibles pour l'inclusion (répondant aux critères d'inclusion) doivent avoir été inclus dans l'étude pendant la période de l'étude (aucune exclusion ou détails sur les raisons des exclusions).	2	1	1
Collecte prospective des données	3	Les données ont été collectées selon un protocole établi avant le début de l'étude.	2	2	2
Critères de jugement appropriés à l'étude	4	Explication sans ambiguïté des critères utilisés pour évaluer le critère principal, qui doit être en accord avec la question posée par l'étude. De plus, les critères de jugement doivent être évalués selon le principe de l'analyse en intention de traiter.	0	2	2
Evaluation impartiale du critère de jugement de l'étude	5	Évaluation en aveugle des critères de jugement objectifs et double-aveugle pour les critères subjectifs. Sinon, les raisons pour lesquelles l'aveuglement n'a pas été utilisé doivent être précisées.	1	1	2
Période de suivi appropriée à l'objectif de l'étude	6	La durée du suivi doit être suffisamment longue pour permettre l'évaluation du critère principal et des éventuels effets indésirables.	1	1	2
Perte de suivi inférieure à 5%	7	Tous les patients doivent être inclus dans le suivi. Sinon, la proportion de perte de suivi ne doit pas dépasser la proportion expérimentant le critère principal.	1	0	0
Calcul prospectif de la taille de l'étude	8	Information sur la taille de la différence détectable d'intérêt avec un calcul de l'intervalle de confiance à 95 %, en fonction de l'incidence attendue de l'événement de résultat, et informations sur le niveau de signification statistique et les estimations de puissance lors de la comparaison des résultats.	2	1	2
CRITÈRES ADDITIONNELS EN CAS D'ÉTUDES COMPARATIVES					
Un groupe de témoin adéquat	9	Disposant d'un test diagnostique ou d'une intervention thérapeutique reconnu comme la meilleure intervention selon les données publiées disponibles.	*	*	2
Groupes contemporains	10	Le groupe témoin et le groupe étudié doivent être gérés pendant la même période (aucune comparaison historique).	*	*	2
Equivalence des groupes de départ	11	Les groupes doivent être similaires concernant les critères autres que ceux étudiés. Absence de facteurs de confusion qui pourraient biaiser l'interprétation des résultats.	*	*	2
Analyses statistiques appropriées	12	Les analyses statistiques sont-elles conformes au type d'étude avec calcul des intervalles de confiance ou du risque relatif.	*	*	2
<i>Note : Les items sont notés 0 (non rapporté), 1 (rapporté mais inadéquat) ou 2 (rapporté et adéquat). Le score global idéal est de 16 pour les études non comparatives et de 24 pour les études comparatives.</i>					
TOTAL			11	10	21

Échelle de Jadad :

Section / Sujet	Item	Description	25
RANDOMISATION		<i>(0 point, si l'essai n'est pas décrit comme randomisé ou si la méthode de randomisation est inappropriée (par exemple, randomisation basée sur la date de naissance, l'ordre d'arrivée, etc.)).</i>	
	1a	L'essai est décrit comme randomisé.	1
	1b	La méthode de randomisation est appropriée (par exemple, l'utilisation d'une table de nombres aléatoires, un générateur de nombres aléatoires par ordinateur).	1
DOUBLE INSU (Double aveugle)		<i>(0 point, si l'essai n'est pas décrit comme en double insu ou si la méthode utilisée pour le double insu est inappropriée (par exemple, si les traitements comparés sont facilement distinguables)).</i>	
	2a	L'essai est décrit comme en double aveugle.	0
	2b	La méthode utilisée pour assurer le double insu est appropriée (par exemple, l'utilisation de placebo identique en apparence, la description des mesures prises pour masquer le traitement des participants et des investigateurs).	*
DESCRIPTION DES ABANDONS ET DES EXCLUSIONS :			
	3a	L'essai décrit le nombre et les raisons des abandons et des exclusions dans chaque groupe.	1
TOTAL / 5			3

VII-5 ANNEXE 5 : Évaluation de la transparence et de la rigueur du rapport de notre revue systématique de la littérature ainsi que de sa qualité méthodologie grâce aux échelles PRISMA et AMSTAR

Échelle PRISMA

	ITEM	RECOMMANDATION	
Titre	1	Identifier le rapport comme une revue systématique, une méta-analyse, ou les deux.	1
Résumé	2	Fournir un résumé structuré incluant, si applicable : contexte ; objectifs ; sources des données ; critères d'éligibilité des études, populations et interventions ; évaluation des études et méthodes de synthèse ; résultats ; limites ; conclusions et impacts des principaux résultats ; numéro d'enregistrement de la revue systématique.	1
INTRODUCTION			
Contexte	3	Justifier la pertinence de la revue par rapport à l'état actuel des connaissances.	1
Objectifs	4	Déclarer explicitement les questions traitées en se référant aux participants, interventions, comparaisons, résultats et à la conception de l'étude (PICOS).	1
MÉTHODES			
Protocole et enregistrement	5	Indiquer si un protocole de revue de la littérature existe, s'il peut être consulté et où (par exemple : adresse web), et, le cas échéant, fournir des informations d'identification, y compris le numéro d'enregistrement.	0
Critères d'éligibilité	6	Spécifier les caractéristiques de l'étude (par exemple : PICOS, durée de suivi) et les caractéristiques du rapport (par exemple : années considérées, langues, statuts de publication) utilisées comme critères d'éligibilité et justifier ce choix.	1
Source d'information	7	Décrire toutes les sources d'information (par exemple : bases de données avec la période couverte, échange avec les auteurs pour identifier des études complémentaires) de recherche et la date de la dernière recherche.	1
Recherche	8	Présenter la stratégie complète de recherche automatisée d'au moins une base de données, y compris les limites décidées, de sorte qu'elle puisse être reproduite.	1
Sélection des études	9	Indiquer le processus de sélection des études (c.-à-d. : triage, éligibilité, inclusion dans la revue systématique, et, le cas échéant, inclusion dans la méta-analyse).	1
Extraction des données	10	Décrire la méthode d'extraction de données contenues dans les rapports (par exemple : formulaires préétablis, librement, en double lecture) et tous les processus d'obtention et de vérification des données auprès des investigateurs.	0
Données	11	Lister et définir toutes les variables pour lesquelles des données ont été recherchées (par exemple : PICOS, sources de financement) et les suppositions et simplifications réalisées.	1
Risque de biais inhérent à chacune des études	12	Décrire les méthodes utilisées pour évaluer le risque de biais de chaque étude (en spécifiant si celui-ci se situe au niveau de l'étude ou du résultat) et comment cette information est utilisée dans la synthèse des données.	0
Quantification des résultats	13	Indiquer les principales métriques de quantification des résultats (par exemple : risk ratio, différence entre les moyennes).	1
Synthèse des résultats	14	Décrire les méthodes de traitement des données et de combinaison des résultats des études, si effectué, y compris les tests d'hétérogénéité (par exemple : I ²) pour chaque méta-analyse.	*

	ITEM	RECOMMANDATION	
Risque de biais transversal aux études	15	Spécifier toute quantification du risque de biais pouvant altérer le niveau de preuve global (par exemple : biais de publication, rapport sélectif au sein des études).	0
Analyses complémentaires	16	Décrire les méthodes des analyses complémentaires (par exemple : analyses de sensibilité ou en sous-groupes, méta-régression), si effectuées, en indiquant celles qui étaient prévues a priori.	*
RÉSULTATS			
Sélection des études	17	Indiquer le nombre d'études triées, examinées en vue de l'éligibilité, et incluses dans la revue, avec les raisons d'exclusion à chaque étape, de préférence sous forme d'un diagramme de flux.	1
Caractéristiques des études sélectionnées	18	Pour chaque étude, présenter les caractéristiques pour lesquelles des données ont été extraites (par exemple : taille de l'étude, PICOS, période de suivi) et fournir les références.	1
Risque de biais relatif aux études	19	Présenter les éléments sur le risque de biais de chaque étude et si possible, toute évaluation des conséquences sur les résultats (voir item 12).	0
Résultats de chaque étude	20	Pour tous les résultats considérés (positifs ou négatifs), présenter, pour chaque étude :	
		a) une brève synthèse des données pour chaque groupe d'intervention ;	1
		b) les amplitudes d'effets estimés et leurs intervalles de confiance, idéalement avec un graphique en forêt (forest plot).	0
Synthèse des résultats	21	Présenter les principaux résultats de chaque méta-analyse réalisée, incluant les intervalles de confiance et les tests d'hétérogénéité.	*
Risque de biais transversal aux études	22	Présenter les résultats de l'évaluation du risque de biais transversal aux études (voir item 15).	0
Analyse complémentaire	23	Le cas échéant, donner les résultats des analyses complémentaires (par exemple : analyses de sensibilité ou en sous-groupes, méta-régression [voir item 16]).	*
DISCUSSION			
Synthèse des niveaux de preuve	24	Résumer les principaux résultats, ainsi que leur niveau de preuve pour chacun des principaux critères de résultat ; examiner leur pertinence selon les publics concernés (par exemple : établissements ou professionnels de santé, usagers et décideurs).	1
Limites	25	Discuter des limites au niveau des études et de leurs résultats (par exemple : risque de biais), ainsi qu'au niveau de la revue (par exemple : récupération incomplète de travaux identifiés, biais de notification).	1
Conclusions	26	Fournir une interprétation générale des résultats dans le contexte des autres connaissances établies et les impacts pour de futures études.	1
Financement	27	Indiquer les sources de financement de la revue systématique et toute autre forme d'aide (par exemple : fourniture de données) ; rôle des financeurs pour la revue systématique.	1
TOTAL /			17

Échelle AMSTAR

	OUI	NON	IMPOSSIBLE DE RÉPONDRE	SANS OBJET	
1. Un plan de recherche établi à priori est-il fourni ?	X				
2. La sélection des études et l'extraction des données ont-elles été confiées à au moins deux personnes ?	X				
3. La recherche documentaire était-elle exhaustive ?	X				
4. La nature de la publication (littérature grise, par exemple) était-elle un critère d'inclusion ?		X			
5. Une liste des études (incluses et exclues) est-elle fournie ?	X				
6. Les caractéristiques des études incluses sont-elles indiquées ?	X				
7. La qualité scientifique des études incluses a-t-elle été évaluée et consignée ?	X				
8. La qualité scientifique des études incluses dans la revue a-t-elle été utilisée de façon adéquate dans la formulation des conclusions ?		X			
9. Les méthodes utilisées pour combiner les résultats des études sont-elles appropriées ?			X		
10. La probabilité d'un biais de publication a-t-elle été évaluée ?		X			
11. Les conflits d'intérêts ont-ils été déclarés ?	X				
TOTAL / 11	7				

AUTEURS : AGUERA Elise et FERRÉ Cécile

TITRE : Revue Systématique de la Littérature sur l'Utilisation de l'Intelligence Artificielle dans le Diagnostic des Mélanomes et des Carcinomes non Mélanocytaires en Soins Primaires.

DIRECTEUR DE THÈSE : Professeur BIREBENT Jordan

LIEU ET DATE DE SOUTENANCE : Faculté de Rangueil – 133 route de Narbonne - 17 octobre 2024

Résumé en français : Les cancers cutanés, mélanocytaires et non mélanocytaires, sont en forte augmentation, dans le monde, en particulier en Australie. La France n'est pas épargnée par ce phénomène, l'incidence de ces cancers ayant triplé entre 1990 et 2023. Le manque de dermatologues entraîne des délais de consultation de plus de 60 jours, contraignant les médecins généralistes à poser des diagnostics, souvent sans formation adéquate. L'intelligence artificielle (IA) a émergé comme une solution pour améliorer le diagnostic et la prise en charge des cancers cutanés.

Notre étude vise à évaluer l'utilisation de l'IA en soins primaires pour aider les médecins généralistes à diagnostiquer les cancers de la peau. Nous avons réalisé une revue systématique de la littérature en suivant les recommandations PRISMA ce qui a permis d'inclure 25 études pertinentes. Parmi elles, 25 % portaient sur les soins primaires et rapportaient des résultats discordants sur la performance diagnostique de l'IA par rapport à celle des médecins généralistes dans un environnement clinique. Cependant, les résultats montraient également que l'IA, notamment multimodale combinant images et métadonnées ou via des réseaux de neurones convolutifs, a une précision diagnostique équivalente ou supérieure à celle des dermatologues dans des environnements contrôlés.

Malgré son potentiel, l'IA nécessite encore des validations cliniques robustes pour une utilisation généralisée, en particulier sur des populations sous-représentées dans les bases de données d'entraînement. Des recherches supplémentaires sont nécessaires pour évaluer son efficacité en dehors des environnements de tests théoriques et son acceptabilité par les médecins et les patients.

Titre et résumé en anglais : Systematic Review of the Literature on the Use of Artificial Intelligence in the Diagnosis of Melanomas and Non-Melanocytic Carcinomas in Primary Care.

Melanocytic and non-melanocytic skin cancers are on the increase worldwide, particularly in Australia. France is not immune to this phenomenon, with the incidence of these cancers tripling between 1990 and 2023. The shortage of dermatologists is leading to appointment delays of over 60 days, forcing general practitioners (GPs) to make diagnoses, often without adequate training. Artificial intelligence (AI) has emerged as a solution to improve the diagnosis and management of skin cancers.

Our study aims to evaluate the use of AI in primary care to help GPs diagnose skin cancers. We conducted a systematic review of the literature following PRISMA recommendations, which resulted in the inclusion of 25 relevant studies. Of these studies, 25% focused on primary care and reported discordant results on the diagnostic performance of AI compared with that of GPs in a clinical setting. However, the results also showed that AI, notably multimodal combining images and metadata or via convolutional neural networks, has a diagnostic accuracy equivalent or superior to that of dermatologists in controlled environments.

Despite its potential, AI still requires robust clinical validations for widespread use, particularly on populations under-represented in training databases. Further research is needed to assess its effectiveness outside theoretical testing environments and its acceptability to physicians and patients.

Mots-Clés : cancer de la peau, mélanome, carcinome épidermoïde, carcinome basocellulaire, intelligence artificielle, apprentissage automatique, apprentissage profond, réseau neuronal convolutif, machine à vecteurs de support, soins primaires, dépistage.

Discipline administrative : MÉDECINE GÉNÉRALE

Faculté de Santé – 37 allées Jules Guesde - 31000 TOULOUSE - France
